

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITE DE MONTPELLIER

En Agronomie

École doctorale [GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau]

UMR AGAP

Identification de gènes impliqués dans la régulation de la mise en place de la paroi secondaire chez le sorgho

Présentée par Lauriane HENNET
Le 12 décembre 2019

Sous la direction de Delphine LUQUET
et Hervé ETIENNE

Devant le jury composé de

M Simon HAWKINS, Professeur, Université de Lille

M Herman HOFTE, Directeur de Recherche, INRA Versailles

Mme Jacqueline GRIMA-PETTENATI, Directrice de Recherche, CNRS Toulouse

M Pascal GANTET, Professeur, Université de Montpellier

Mme Delphine LUQUET, Directrice de Recherche CIRAD

Rapporteur

Rapporteur

Examinatrice

Président du Jury

Directrice de Thèse



UNIVERSITÉ
DE MONTPELLIER



Résumé

Les parois secondaires (PS) des plantes représentent la majorité de la biomasse renouvelable et participent aux fonctions agronomiques des plantes comme le transport des nutriments, la résistance à la verse et aux stress. Le sorgho est une graminée en C4 bien adaptée aux conditions limitantes en nutriments et en eau. L'amélioration des connaissances relatives à la mise en place des PS chez le sorgho a pour objectif de contribuer au développement de variétés adaptées aux filières de valorisation de la biomasse incluant l'alimentation animale, la production énergétique et le développement de bio-matériaux. Les objectifs de la thèse présentée ici sont de contribuer à la compréhension des mécanismes moléculaires de la mise en place des PS par l'identification des réseaux de gènes impliqués et la validation de leur rôle. Dans un premier temps, une analyse phylogénétique comprenant 9 espèces, nous a permis d'identifier les liens d'homologies au sein des familles de facteurs de transcription (FT) NAC et MYB, qui sont les principaux régulateurs de la mise en place des PS chez *Arabidopsis* et d'autres espèces. Nous avons ensuite construit les réseaux de co-expression géniques impliqués dans la mise en place des entre-nœuds en développement du sorgho et nous nous sommes plus particulièrement intéressés aux réseaux comprenant des NAC et des MYB ainsi qu'à ceux présentant des enrichissements fonctionnels liés aux PS. Puis les gènes identifiés par ces approches ont été confrontés à ceux situés dans les zones chromosomiques impactant la variabilité de la composition des PS. Dans un deuxième temps, nous avons développé des approches de transformation génétique afin de caractériser le rôle des gènes mis en évidence par génétique inverse. Nous avons identifié 6 modules de gènes co-exprimés enrichis en gènes et en ontologies géniques liés aux parois et en particulier 2 modules dont les gènes sont potentiellement impliqués dans la mise en place des PS. En outre, 122 NAC et 135 MYB ont été identifiés chez le sorgho. Parmi eux, 14 NAC et 25 MYB sont impliqués dans les 6 réseaux de co-expression de gènes fortement enrichis en gènes de paroi. Parmi eux, 19 sont des homologues de gènes déjà validés chez d'autres espèces pour leur rôle dans la mise en place des PS et 20 n'ont pas de rôle connu. La comparaison des gènes faisant partie des 6 modules impliqués dans la mise en place des parois et des régions chromosomiques associées à la composition de la biomasse a révélé un intersect de 70 gènes et quelques gènes de structure impliqués dans la biosynthèse des parois mais peu de convergence concernant les FT NAC et MYB. Bien que l'optimisation des protocoles de transformation génétique stable n'ait pas abouti, un protocole efficace de transformation transitoire basé sur des protoplastes a été optimisé. Cet outil sera notamment mobilisé pour tester les rôles des FT identifiés dans les réseaux enrichis en gènes de parois.

Abstract

Plants Secondary Cell Walls (SCW) represent most of the renewable biomass and participate in the agronomic functions of plants such as nutrient transport, stress resistance and stiffness. Sorghum is a C4 grass well adapted to nutrient and water limiting conditions. Improving knowledge SCW establishment in sorghum aims at contributing to the development of varieties adapted to biomass sector including animal feed, energy production and the development of bio-materials. The objectives of the thesis presented here are to contribute to the understanding of the molecular mechanisms of SCW regulation and biosynthesis by identifying genes networks involved in this process and validating their role. As a first step, a phylogenetic analysis including 9 species, allowed us to identify homologies within families of NAC and MYB TF, which are the main regulators of SCW establishment in *Arabidopsis* and other species. We then built gene co-expression networks of sorghum developing internode and focused on NAC and MYB networks as well as on networks with functional enrichment related to SCW. Then, genes identified by these approaches were compared to those located in the chromosomal regions impacting the variability of SCW composition. In a second step we have developed genetic transformation approaches to characterize the role of identified genes by reverse genetics. We have identified 6 modules of co-expressed genes enriched with genes and gene ontologies related to cell walls and in particular 2 modules whose genes are potentially involved in the establishment of SCW. In addition, 122 NAC and 135 MYB were identified in sorghum. Among them, 14 NAC and 25 MYB are involved in the 6 co-expression networks highly enriched in cell wall genes. Of these, 19 are homologues of genes already validated in other species for their role SCW establishment and 20 have no known role. The comparison of the genes belonging to the 6 modules involved in cell wall establishment and chromosomal regions associated with the composition of the biomass revealed an intersect of 70 genes and some structural genes involved in the biosynthesis of the walls but little convergence concerning NAC and MYB TF. Although the optimization of stable genetic transformation protocols was unsuccessful, an efficient protoplast-based transient transformation protocol was optimized. This tool will be mobilized to test the roles of identified TF in networks enriched with cell wall genes.

Remerciements

Tout d'abord, je souhaite remercier mes deux rapporteurs, Herman Höfte et Simon Hawkins, d'avoir accepté d'évaluer mes travaux de thèse. Je remercie également Jacqueline Grima-Pettenati et Pascal Gantet d'avoir accepté de participer à mon jury de thèse.

Je remercie Matthieu Reymond, Bertrand Muller, Patrice This, Jacqueline Grima-Pettenati et Emmanuel Guiderdoni d'avoir accepté de faire partie de mon comité de suivi de thèse.

Je remercie Delphine Luquet et Hervé Etienne de m'avoir permis de réaliser cette thèse en assurant la direction.

Je remercie David Pot et Nancy Terrier pour leur encadrement, leurs enseignements et leurs conseils au cours de ces trois années de thèse éprouvantes mais enrichissantes.

Je remercie Frédérique Richaud, Angélique Berger, Caroline Calatayud et Maelle Rios pour leur immense contribution à ce travail sous la forme, notamment, d'innombrables heures en chambre de culture et sous la hotte.

Je remercie Marion Dupouy et Aurélien Cottin pour leur inestimable aide pour mes analyses bioinformatiques.

Je remercie Ian Godwin et Karen Massel pour leur accueil et leur supervision pendant mon séjour dans leur laboratoire.

Je remercie les doctorants d'AGAP (et autre) qui ont fait de ces trois années une belle expérience ainsi que l'ensemble des personnes avec lesquelles j'ai pu échanger à propos de ma thèse ou de tout autre sujet.

Enfin, je remercie Hugo et mes amis d'être toujours à mon écoute que ce soit pour des idées saugrenues ou des doutes. Je remercie également ma famille pour son soutien sans lequel il n'aurait pas été possible de réaliser de si longues études dans de si bonnes conditions.

Lauriane

Table des matières

I. Introduction générale.....	9
II. Chapitre I : Synthèse bibliographique introductive	13
1. Contexte	15
2. Le sorgho, classification et usages agronomiques	16
3. Atouts agronomiques et génétiques du sorgho	17
4. Morphologie du sorgho	20
5. Cycle de développement	27
6. Structure de la tige et organisation des parois	28
7. Description des composés pariétaux	35
7.1. Cellulose	35
7.2. Hémicelluloses.....	37
7.3. Lignines	41
7.4. Pectines	45
7.5. Protéines pariétales	47
7.6. Flavonoïdes.....	47
8. Composition générale des parois	51
9. Biosynthèse des principaux composants des parois secondaires	57
9.1. Biosynthèse de la cellulose.....	57
9.2. Biosynthèse des hémicelluloses.....	58
9.3. Biosynthèse des lignines.....	58
10. Variation de la composition des parois en réponse aux « stress »	61
10.1. Stress abiotiques	61
10.2. Stress biotiques	63
11. Déterminisme génétique et évaluation de la qualité de la biomasse	64
11.1. Evaluation de la qualité de la biomasse.....	66
11.2. Déterminisme génétique de la qualité de la biomasse	70
12. Facteurs de transcription impliqués dans la régulation des parois secondaires .	72
12.1. Facteurs de Transcription NAC	73
12.2. Facteurs de transcription MYB.....	79
12.3. Autres régulateurs des parois secondaires	84
13. Objectifs de la thèse et organisation du manuscrit	86
III. Chapitre II : Identification de gènes candidats impliqués dans la détermination de la qualité de la biomasse chez le sorgho	87
1. Introduction.....	88
1.1. Analyses phylogénétiques	89

1.2.	Analyse de réseaux de co-expression de gènes	93
1.3.	Génétique d'association.....	98
2.	Matériels et méthodes	102
2.1.	Analyses phylogénétiques des familles MYB et NAC.....	102
2.2.	Réseaux de co-expression.....	103
3.	Résultats.....	112
3.1.	Phylogénie des familles NAC et MYB : apport pour l'identification des FT régulant la mise en place des parois secondaires chez le sorgho	112
3.2.	Identification d'acteurs de la biosynthèse des parois par la construction de réseaux de co-expression de gènes.....	123
3.3.	Gènes mis en évidence par des analyses de génétique d'association	146
4.	Discussion	149
4.1.	Mise à jour des classifications phylogénétique des FT MYB et NAC dans l'objectif de faciliter la compréhension des acteurs principaux de la mise en place des parois secondaires.	149
4.2.	Identification de réseaux de co-expression de gènes contribuant au développement des entre-nœuds	156
4.3.	Evaluation des liens d'homologies entre les NAC et les MYB par l'analyse des réseaux de co-expression	158
4.4.	Gènes associés à la qualité de la biomasse en lien avec les FT NAC et MYB et les réseaux de co-expression liés à la mise en place des parois	164
5.	Conclusion	168
IV.	Chapitre III : Utilisation de la transformation génétique pour la validation fonctionnelle de gènes impliqués dans la mise en place des parois secondaires.....	170
1.	Introduction.....	171
1.1.	Principe et intérêt de la transformation génétique	171
1.2.	Principe de transfert de l'ADN dans le matériel végétal	173
1.3.	Constitution des constructions et principe de la stratégie CRISPR/Cas9.....	181
1.4.	Etat des lieux de la transformation stable chez les graminées et le sorgho ..	185
1.5.	Objectifs généraux des essais de transformation stable.....	188
1.6.	Choix des cibles génétiques pour la transformation stable.....	188
1.7.	Etat de l'art la production et la transformation des protoplastes	193
2.	Matériels et méthodes	195
2.1.	Construction des plasmides	195
2.2.	Transformation stable par biolistique et <i>Agrobacterium tumefaciens</i>	204
3.	Résultats.....	215
3.1.	Essais de transformation stable.....	215

3.2.	Optimisation de la production et de la transformation de protoplastes de sorgho	221
4.	Discussion	237
4.1.	Essais de transformation stable	237
4.2.	Production et transformation de protoplastes	243
V.	Discussion générale	247
1.	Rappel des objectifs et des stratégies	248
2.	Synthèse des résultats	250
2.1.	Analyse de la phylogénie des NAC et des MYB	250
2.2.	Analyse des réseaux de co-expression de gènes	250
2.3.	Mise en évidence de régions génomiques associées à la variabilité de la biomasse par analyse de génétique d'association	252
2.4.	La transformation stable chez le sorgho reste un challenge	252
2.5.	La transformation transitoire sur protoplaste de sorgho est une voie rapide de caractérisation de gènes	253
3.	Arabidopsis, un modèle absolu ?	253
4.	Découverte et caractérisation de gènes chez les espèces non modèles	257
5.	Les difficultés de la transformation génétique	260
VI.	Contributions	264
1.	Contributions au chapitre II	264
2.	Contributions au chapitre III	264
VII.	Références	265
VIII.	Annexes	324

Introduction générale

Le sorgho est la 5ème céréale mondiale en termes de production de grains derrière le blé, l'orge, le maïs et le riz. Au niveau mondial, une demi-milliard de personnes dépendent du sorgho quotidiennement et consomment le grain et la farine de sorgho comme alimentation de base (Frère *et al.*, 2011; Paterson, 2009). En plus de son importance pour la nutrition humaine, les parties végétatives et les grains du sorgho sont également utilisés pour la production de fourrage, d'énergie et de matériaux renforcés en fibres végétales (Anami *et al.*, 2015; Antonopoulou *et al.*, 2008; Barcelos *et al.*, 2016; Brenton *et al.*, 2016a; Carpita et McCann, 2008; Carvalho et Rooney, 2017; Dien *et al.*, 2009; Mullet *et al.*, 2014; Thomas *et al.*, 2019).

Le sorgho est une céréale en C4 capable de fixer efficacement le carbone et de produire d'importantes quantités de biomasse végétative dans une grande variété d'agrosystèmes. Grâce à sa tolérance au manque d'intrants, à la variabilité des températures et au stress hydrique (Brown *et al.*, 2008; Rooney *et al.*, 2007; Sanderson *et al.*, 1992; Schmidhuber et Tubiello, 2007; Zegada-Lizarazu *et al.*, 2012), le sorgho s'avère un bon candidat pour produire la biomasse requise pour la production d'énergie et de matériaux à partir de bio-ressources. De plus, le sorgho présente plusieurs caractères de plante modèle (Aitken *et al.*, 2014; Mullet *et al.*, 2014; Swaminathan *et al.*, 2010). Son génome est relativement simple comparé à d'autres graminées avec une taille de 780Mb et n'a pas subi de nombreuses duplications récentes (Paterson *et al.*, 2009; Schnable *et al.*, 2011). Le sorgho présente également une grande diversité génétique (Billot *et al.*, 2013a; Morris *et al.*, 2013; Sekhon *et al.*, 2016) et phénotypique qui sont mises à contribution dans les programmes de sélections en milieu tempéré et tropical. Enfin, la transformation génétique au moyen de vecteurs biologiques et physiques et des outils d'édition ont déjà été utilisés dans le cadre de l'étude du sorgho. La répétabilité des protocoles reste à optimiser mais des taux de transformation jusqu'à 40% ont été obtenus (Che *et al.*, 2018; Liu et Godwin, 2012; Wu *et al.*, 2014).

Chez le sorgho, la majorité de la biomasse des parties aériennes est allouée à la tige sous la forme de sucres solubles et de parois cellulaires. Dans cette thèse nous nous intéresserons majoritairement à la mise en place des parois secondaires qui comptent pour une très grande part de la biomasse et sont la source principale pour la production d'énergie, de matériaux

renforcés par des fibres végétales, et de fourrages. Les parois secondaires du sorgho sont composées d'environ 50% de cellulose, 43% d'hémicellulose, des polysaccharides également présents dans les parois primaires, et 7% de lignines (Trouche *et al.*, 2014a), un polymère phénolique spécifique des parois secondaires. Ces différents éléments pariétaux s'accumulent autour de cellules dont la croissance est terminée. Dans la tige, il s'agit des cellules qui forment des vaisseaux conducteurs et les cellules sclérenchymateuses qui assurent le soutien des structures de la tige. De nombreux caractères d'adaptation et d'intérêt agronomique reposent sur la composition des parois secondaires de la tige comme la structure, le transport de la sève et la résistance aux stress. En plus de ces considérations générales, les parois secondaires sont importantes dans le cadre des utilisations où la tige est la matière première. Leur composition joue donc un rôle clé dans la qualité agronomique du sorgho, les utilisations fourragères et la transformation industrielle. Cependant, les sélectionneurs ne parviennent pas à développer facilement de nouvelles variétés qui rassemblent les traits désirés, et parfois antagonistes, comme la résistance à la verse, le contenu en polysaccharides et la digestibilité de la tige. Comprendre les mécanismes de la formation des parois secondaires contribuerait à l'amélioration de l'efficacité de la sélection dans l'objectif de parvenir à satisfaire les besoins des agriculteurs et des consommateurs.

Chez le sorgho, des avancées significatives ont été permises par l'analyse de mutants naturels ou induits. Les gènes responsables du phénotype de « nervure brune » (Bmr) qui est accompagné d'une meilleure dégradabilité des parois ont été étudiés en premier. Huit gènes responsables de ce phénotype ont été découverts chez le sorgho (Saballos *et al.*, 2012; Sattler *et al.*, 2014). Trois d'entre eux ont été caractérisés et correspondent à des enzymes de la voie de la biosynthèse des lignines. Deux ont été utilisés pour développer des variétés à destination de l'alimentation animale (Pedersen *et al.*, 2008). En plus du phénotype Bmr, d'autres mutants liés à la composition des parois secondaires ont été identifiés. En effet, Petti *et al.* (Petti *et al.*, 2013, 2015) ont identifié les mutants REDforGREEN et dwarf1.1 dont les compositions en lignines et en cellulose sont affectées dans la tige et les feuilles.

Si des analyses de populations biparentales (Murray *et al.*, 2008a, 2008b; Shiringani et Friedt, 2011) et de panels de diversités larges (Brenton *et al.*, 2016a; Li *et al.*, 2018b) ont permis l'identification de régions génomiques potentiellement impliquées dans la variabilité de la composition des parois secondaires, ces approches ont échoué à apporter une compréhension profonde du contrôle génétique de ce caractère. En parallèle, différentes études de transcriptomique ayant l'ambition de contribuer à la compréhension des mécanismes de mise

en place des parois secondaires ont mis en évidence les patrons d'expression différentiels de listes de gènes dans des entre-nœuds en développement (Kebrom *et al.*, 2017; McKinley *et al.*, 2016; Rai *et al.*, 2016; Shakoore *et al.*, 2014). Cependant, une faible concordance des acteurs révélés entre les approches génétiques (QTL/QTN) et génomiques (transcriptomique) a été observée jusqu'à présent. Seuls quelques gènes candidats de structure impliqués dans la biosynthèse des parois ont été identifiés et des travaux supplémentaires sont nécessaires. En particulier, les facteurs de transcription (FT) sont des régulateurs fins des processus biologiques et peu de travaux s'y sont intéressés chez le sorgho dans le contexte de la mise en place des parois secondaires.

Les travaux chez *Arabidopsis* ont fourni des bases solides concernant les FT impliqués dans la régulation de la mise en place des parois secondaires. La plupart de ces gènes font partie de la famille des FT MYB (Myb proto-oncogene like) et des NAC (NAM, ATAF et CUC2) (Zhong *et al.*, 2008). Certains sont des régulateurs principaux et généraux de la mise en place des parois secondaires tandis que d'autres agissent à des niveaux plus en aval en régulant des voies spécifiques de biosynthèse et d'organisation des composants pariétaux. Il a été démontré que la fonction de certains des régulateurs principaux de ces familles ont une fonction conservée, au moins en partie, chez le riz, le maïs, le peuplier et eucalyptus (Goicoechea *et al.*, 2005; Zhong *et al.*, 2011, 2013). Le socle de connaissances développé concernant ces FT chez *Arabidopsis* et quelques autres plantes est une opportunité pour accélérer et faciliter la découverte de gènes impliqués dans la régulation des parois secondaires chez le sorgho et les graminées. Chez le sorgho, la régulation de la mise en place des parois secondaires est très mal connue avec un seul facteur de transcription (SbMYB60) capable d'induire la voie de biosynthèse des lignines qui a été validé récemment (Scully *et al.*, 2016a).

Dans l'objectif de contribuer à la compréhension des mécanismes moléculaires de la mise en place des parois, les réseaux de FT et de gènes structuraux impliqués dans ce processus ont été explorés. Dans un premier temps nous avons cherché à identifier les gènes impliqués dans la mise en place des parois secondaires chez le sorgho en tirant partie des informations existantes chez *Arabidopsis* et d'autres espèces et en les associant à des analyses génomiques et génétiques développées chez le sorgho. Ainsi, nous nous sommes tout d'abord focalisés sur les FT des familles MYB et NAC en menant une étude de génomique comparative permettant de mettre en évidence les homologies existantes entre les FT du sorgho et les gènes connus chez d'autres espèces pour leur rôle dans la régulation des parois secondaires. Ensuite, des données transcriptomiques basées sur des dynamiques développementales d'entre-nœuds ont

été utilisées pour construire des réseaux de co-expression dans l'objectif de révéler les réseaux de régulation et les FT impliqués dans la mise en place des parois secondaires. Enfin, les gènes candidats identifiés par les approches précédentes ont été comparés aux résultats obtenus par une analyse de génétique de la diversité globale visant à identifier les régions impliquées dans la variabilité de la composition des parois secondaires. Dans un deuxième temps, nous avons cherché à renforcer les connaissances acquises en première partie en caractérisant les gènes identifiés via des approches de génétique inverse. Ces essais ont été l'occasion de chercher à maîtriser la transformation stable chez le sorgho et de développer un protocole de transformation transitoire rapide afin de disposer d'outils efficace de caractérisation du rôle des gènes.

Chapitre I : Synthèse bibliographique introductive

1. Contexte

Cette thèse s'inscrit dans un contexte de changement climatique, de mondialisation et d'une population humaine grandissante, qui sont quelques-unes des problématiques autour desquelles s'articulent les enjeux du XXIème siècle. La recherche de la réduction des émissions de gaz à effets de serre responsables du réchauffement climatique en est un premier volet. L'augmentation de l'efficacité des moyens de production d'énergie, la captation des gaz à effet de serre et la sobriété énergétique nécessitent des efforts pour la recherche de combustibles renouvelables. Les agro-carburants de première génération qui reposent sur l'utilisation du grain et des sucres solubles produits par les végétaux ont subi un essor au cours des décennies passées mais restent non consensuelles car elles accaparent des ressources et des terres arables pour la production d'une biomasse dont l'usage est en concurrence avec la consommation humaine (Mohr et Raman, 2013). En effet, les céréales utilisées dans ce cadre demandent des conditions de culture exigeantes et sont exclusivement destinées à ces productions. Les conséquences de la production des agro-carburants de première génération s'étendent jusqu'aux problématiques de déforestation et de marché mondial des matières premières. Il y a donc une nécessité de développer des variétés et des systèmes de production multi-usages dont les exigences sont moindres. Les agro-carburants de deuxième génération répondent à une partie de ces contraintes par l'utilisation de la biomasse de la plante entière.

De la même manière que pour les carburants, les ressources fossiles non renouvelables sont à la base de la fabrication d'une grande majorité des matériaux. Les matériaux renforcés par des fibres végétales sont de bons candidats pour remplacer les plastiques et bétons et l'utilisation de la biomasse peut s'étendre au développement de nouveaux types de tissus, de produits chimiques, de molécules plateformes bio-sourcées servant de base à la synthèse d'autres molécules (Furtwengler et Avérous, 2018; Jawaïd *et al.*, 2017). La stratégie de bioraffinerie permet d'envisager la valorisation optimale des composants des agro-ressources.

Le deuxième volet des enjeux du changement climatique est la prise en compte de conditions de culture contraintes et modifiées. En effet, le changement climatique a pour conséquence une augmentation globale des températures, la diminution des ressources en eau et des événements climatiques extrêmes de plus en plus fréquents (Thornton *et al.*, 2018). La tendance des dernières décennies a été à l'uniformisation des pratiques culturales, à la sélection de plantes très productives peu adaptées à des conditions variables de croissance et

dont la productivité repose sur des ressources externes que sont l'irrigation, les engrais et pesticides utilisés en grande quantité. De plus, dans les pays tropicaux où l'agriculture doit s'intensifier, des variétés adaptées à ces contextes sont donc nécessaires. Que ce soit dans les pays tempérés ou tropicaux, il y a donc un besoin de nouvelles variétés adaptées aux contextes locaux et variables qui sauront répondre aux évolutions climatiques du futur.

2. Le sorgho, classification et usages agronomiques

Le sorgho commun, *Sorghum bicolor*, est une plante appartenant au clade des monocotylédones et à la famille des poacées (graminées). C'est une espèce très proche de la canne à sucre, du maïs et du miscanthus. *Sorghum biocolor* ssp. *bicolor*, le sorgho cultivé, présente 5 races (guinea, durra, bicolor, caudatum, kafir) en plus de la classification fonctionnelle (sorgho grain, fourrager, sucré, biomasse) qui désigne l'usage agronomique préférentiel de la variété et qui est inter-raciale.

Le sorgho est cultivé dans les régions arides et semi arides de l'Afrique et, s'il a été domestiqué en Ethiopie en ~3000 AC (FAOSTAT, 2001), sa culture s'est largement répandue en Afrique puis au Moyen Orient, en Inde et jusqu'en Chine (Winchell *et al.*, 2018). Ces dernières décennies, la culture du sorgho se développe partout dans le monde, notamment en Russie, en Europe, en Amérique du Sud, en Australie et aux Etats-Unis. Le sorgho sous forme de grain est la 5^{ème} céréale la plus produite du monde (en 2017, 57.6 millions de tonnes de sorgho ont été produites), la 3^{ème} aux USA (après le blé et le maïs) et la plus importante céréale d'été en Australie où le sorgho est plus cultivé que le maïs (FAOSTAT 2017). Les plus grands producteurs de sorgho sont les USA, le Nigeria, le Mexique, l'Ethiopie, l'Inde et le Soudan (au-delà de 3M de T produites), la France étant le 23^{ème} producteur mondial et le premier producteur d'Europe (FAOSTAT 2017).

Les utilisations du sorgho sont variées, différents idéotypes de sorgho sont donc cultivés selon les utilisations visées. Le sorgho est l'aliment de base pour des centaines de millions de personnes dans le monde (Frère *et al.*, 2011) qui utilisent, au-delà du fourrage pour le bétail, les grains entiers ou concassés pour de nombreuses préparations traditionnelles, pains, bouillies et biscuits entre autres. En dehors des pays Africains et de l'Inde, grains et tiges sont également utilisés en alimentation animale aux Etats Unis, en Europe et en Amérique du sud. Le sorgho suscite également de forts intérêts pour la production d'agro-carburants de deuxième génération et de matériaux renforcés par des fibres végétales. Cet intérêt récent pour le sorgho émerge face aux enjeux du changement climatique et aux attentes actuelles

d'une agriculture moins dépendante des intrants compte tenu de ses atouts agronomiques présentés dans la section suivante.

3. Atouts agronomiques et génétiques du sorgho

Le sorgho est une céréale aux nombreux atouts agronomiques. Elle se développe également en tant que plante modèle grâce à ses caractéristiques génétiques.

La culture du sorgho est possible dans une grande variété d'environnements, en particulier sur des terres arides ou semi-arides en conditions de hautes températures où la culture d'autres céréales comme la canne à sucre ou le maïs est difficile (Rooney *et al.*, 2007). En comparaison avec le maïs, certaines variétés de sorgho ont le potentiel de produire plus de biomasse totale en particulier en conditions de sécheresse (Hallam *et al.*, 2001). Même si le maïs est plus performant en conditions optimales d'irrigation, le gain de rendement du sorgho dans des conditions de stress hydrique par rapport au maïs est important, de par ses capacités de résistance à la sécheresse ainsi que sa capacité à entrer en dormance pendant ces périodes (Norwood et Currie, 1997; Poehlman, 1987; Propheter *et al.*, 2010; Rocateli *et al.*, 2012; Rooney *et al.*, 2007; Smith *et al.*, 1987). De surcroît, le sorgho ne requière qu'1/3 des besoins en eaux de la canne à sucre et 1/2 de ceux du maïs (Enciso *et al.*, 2015; Farré et Faci, 2006; Garofalo et Rinaldi, 2013). De plus, au-delà de son métabolisme en C4, le sorgho résiste bien aux conditions de manque d'eau et de nutriment grâce à sa bonne récupération des nutriments du sol et son utilisation efficace de l'azote (Gardner *et al.*, 1994). La capacité du sorgho à résister à la sécheresse est aussi assurée par son réseau extensif de racines, la capacité à enrouler ses feuilles ainsi que la capacité des stomates à récupérer rapidement après le regain de turgescence des feuilles (Poehlman, 1987; Rosenow *et al.*, 1983)

Les atouts agronomiques du sorgho mentionnés ci-dessus attirent de plus en plus d'efforts de recherche. Ainsi, les ressources génétiques se développent autour de cette plante qui a le potentiel de devenir un modèle biologique des graminées en C4.

Le sorgho est une plante diploïde à 10 paires de chromosomes qui rassemblent environ 780 millions de paires de bases (Mb) et 38 000 gènes (McCormick *et al.*, 2018; Paterson *et al.*, 2009). Le sorgho fut la deuxième céréale dont le génome a été entièrement séquencé (après le riz (Sasaki et Project, 2005) et juste avant le maïs (Schnable *et al.*, 2009). En comparaison avec d'autres céréales comme le maïs (2,3 Gb (Schnable *et al.*, 2009) et la canne à sucre (10Gb (Garsmeur *et al.*, 2018)), le génome du sorgho est petit. De surcroît, il est relativement simple n'ayant pas subi de duplication récente (Paterson *et al.*, 2009). Il a même servi de base,

grâce à la proximité des deux espèces, au séquençage du génome très complexe de la canne à sucre (Figure I-1).

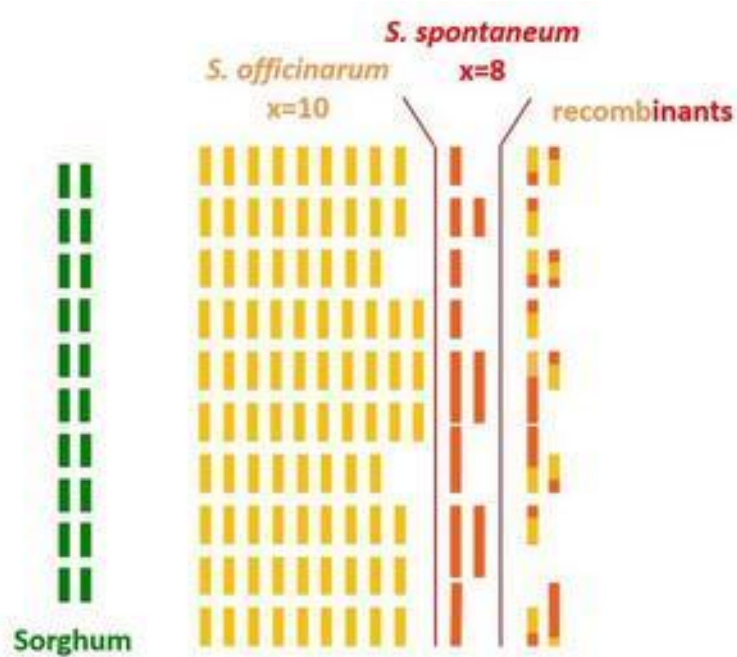


Figure I-1 : Comparaison de la structure du génome du sorgho et des génomes complexes de deux espèces de canne à sucre (*S. officinarum* et *S. spontaneum*) et d'un cultivar moderne typique de canne à sucre (« recombiants »). Chaque barre représente un chromosome, en vert le sorgho, en orange et en rouge deux espèces de canne à sucre dont sont issus les cultivars modernes. Chaque rangée représente des chromosomes homologues. (D'après Garsmeur *et al.*, 2018)

Les outils de génétique sont moins développés chez le sorgho que chez d'autres espèces comme *Arabidopsis* ou le riz. Cependant, grâce à ses caractéristiques génétiques, le sorgho est un modèle attractif pour la recherche en génomique fonctionnelle qui est plus difficile chez le maïs, la canne à sucre, *miscanthus* et d'autres céréales en C4 aux génomes complexes. (Hughes *et al.*, 2014; Paterson *et al.*, 2009; Schnable *et al.*, 2009; Swigonova, 2004). Des outils se développent, comme l'atlas transcriptomique en libre accès MOROKOSHI (<http://sorghum.riken.jp/morokoshi/Home.html>) dont les données sont issues de différentes études indépendantes (Davidson *et al.*, 2012; Dugas *et al.*, 2011; Gelli *et al.*, 2014; Shimada *et al.*, 2015; Yazawa *et al.*, 2013). Une synthèse des régions chromosomiques impliquées dans le contrôle de la variabilité phénotypiques de caractères d'intérêt est disponible (<https://aussorgm.org.au/sorghum-qt1-atlas/>) ainsi qu'une collection de mutants chimiques EMS (Ethyl-Methane Sulfonate) qui a été produite par Zanghuo et son équipe (Jiao *et al.*, 2016). Les mutations ont été induites dans le génome complet de façon non contrôlée sur des graines et les génomes de 256 plantes de la troisième génération ont été séquencés de façon à déterminer les loci mutés. Cette méthode a permis d'introduire des mutations de tout type dans plus de 95% des gènes du sorgho (Jiao *et al.*, 2016). Cette banque de mutants est prometteuse grâce à la faible redondance génétique du sorgho et la large gamme de phénotypes que cette approche génère (Jiao *et al.*, 2016; Xin *et al.*, 2008).

Les approches de génétique inverse se développent également de plus en plus grâce aux techniques de transformation génétique du sorgho qui se perfectionnent (Belide *et al.*, 2017; Che *et al.*, 2018; Liu et Godwin, 2012; Wu *et al.*, 2014). La plupart des techniques ont longtemps eu des taux de transformation atteignant au maximum 5% (Howe *et al.*, 2006; Jambagi *et al.*, 2010), mais certaines équipes parviennent maintenant à des taux de transformation de 15 à 50% (Belide *et al.*, 2017; Do *et al.*, 2016a; Liu et Godwin, 2012; Wu *et al.*, 2014). Malgré ces succès, la transformation du sorgho est encore difficile et les études publiées ces dernières années ont toujours régulièrement des taux de transformation inférieurs à 5%. Les techniques de transformation du sorgho seront développées plus en détail dans le chapitre III.

4. Morphologie du sorgho

La morphologie d'un plant de sorgho mûre est proche de celle des autres graminées. Il comporte une tige principale et un certain nombre de tiges secondaires qui ramifient à partir de la base de la plante et appelées talles. Toutes les tiges sont composées d'une succession

d'unités morphologiques identiques, les phytomères. Un phytomère est constitué d'une feuille, un nœud portant un bourgeon axillaire et un entre-nœud développé en dessous du nœud (Figure I-2A et B)

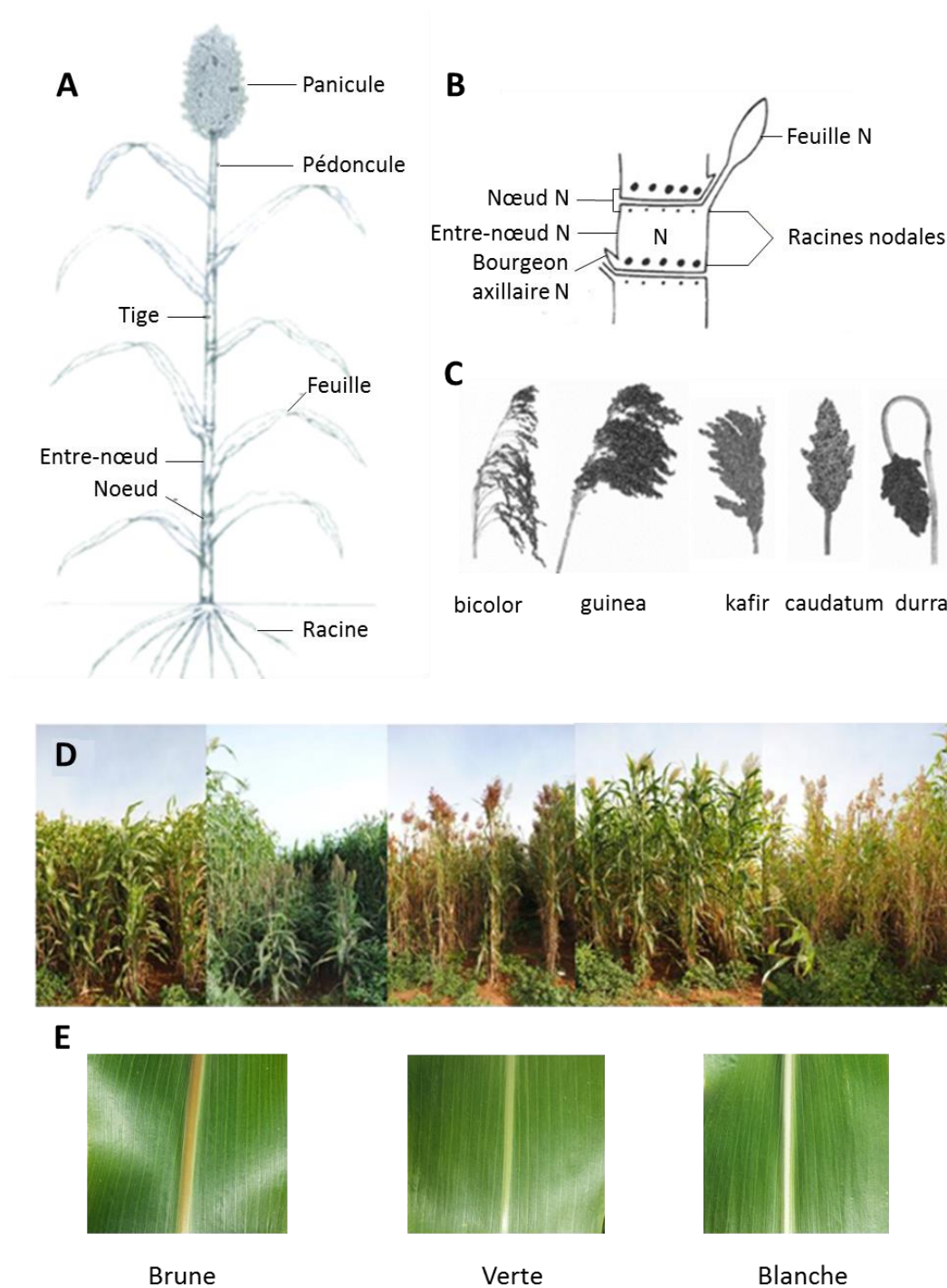


Figure I-2: Diversité et description de l'appareil végétatif du sorgho. Représentation de la morphologie (A) d'une plante de sorgho, (B) d'un phytomère et (C) de différents types de panicules représentatifs des 5 races de sorgho. (D) Aperçu de la diversité morphologique disponible dans le contexte de l'analyse de la production de biomasse aérienne. (E) Différentes couleurs de nervure centrale. (D'après Clerget *et al.*, 2004; Perrier, 2017a; Trouche *et al.*, 2014 et images de nervures de F. De Bellis)

Les phytomères les plus bas peuvent aussi développer d'épaisses racines adventives qui ancrent la plante solidement dans le sol. Les phytomères sont émis successivement par le méristème à l'apex de la tige. Une fois le développement des entre-nœuds achevé, le méristème marque une transition et émet une inflorescence finale. La croissance du sorgho est donc déterminée et le nombre de phytomères est fixé à l'apparition d'une panicule (Figure I-3).

Le système racinaire du sorgho est très développé et croît rapidement et profondément dans le sol sur 1 à 1,5m. Les racines adventives croissent dans un deuxième temps et permettent la formation d'un réseau étendu qui assure une excellente captation de l'eau du sol. C'est le système racinaire du sorgho qui explique en grande partie sa résistance à la sécheresse. A maturité, il peut représenter 20% de la biomasse totale (Chantereau *et al.*, 2013).

La hauteur de la tige est variable selon les variétés de sorgho allant de 50 cm à plus de 5 m (Figure I-2D). La longueur d'un entre-nœud au sein d'un génotype est relativement stable hormis pour les entrenœuds les plus basaux et âgés, qui sont très courts. La longueur des entre-nœuds varie de 5 à 25 cm en fonction des variétés et est majoritairement expliquée par 4 gènes (dits gènes de nanisme)(Yamaguchi *et al.*, 2016). Le nombre d'entrenœuds est, lui, déterminé par la durée du cycle de vie qui dépend entre autres de la sensibilité à la longueur du jour et la période de culture (Chantereau *et al.*, 2013). La présence de talles basaux dépend également de la variété et des facteurs environnementaux. Cependant, seuls un à deux talles sont généralement observés (Chantereau *et al.*, 2013).

Les feuilles s'insèrent au niveau des nœuds à raison d'une feuille par entre-nœud. Elles sont de forme lancéolée et constituées d'une gaine qui entoure la tige sur plusieurs dizaines de centimètres puis d'une ligule qui forme la jonction avec le limbe. Le nombre de feuilles (de 15 à 30), comme le nombre d'entrenœuds est corrélé à la durée de la phase végétative (Figure I-3) et est donc variable selon les variétés. Les feuilles sont les plus larges en leur milieu (maximum de 13cm) et s'étendent en longueur de 30 à 135cm. La couleur de la nervure des feuilles varie selon les variétés et est verte chez les variétés juteuses, blanche chez les variétés dont la tige est peu concentrée en sucres solubles et brune chez les variétés Bmr (Brown Midrib), rares et à faible teneur en lignines (Figure I-2E). La dernière feuille à apparaître avant l'émergence de l'inflorescence marque une transition vers la phase reproductive et porte le nom de feuille drapeau (Figure I-3).

L'inflorescence s'insère sur le rachis de la tige (partie terminale) et est ramifiée dès sa base en axes primaires puis secondaires voire tertiaires.

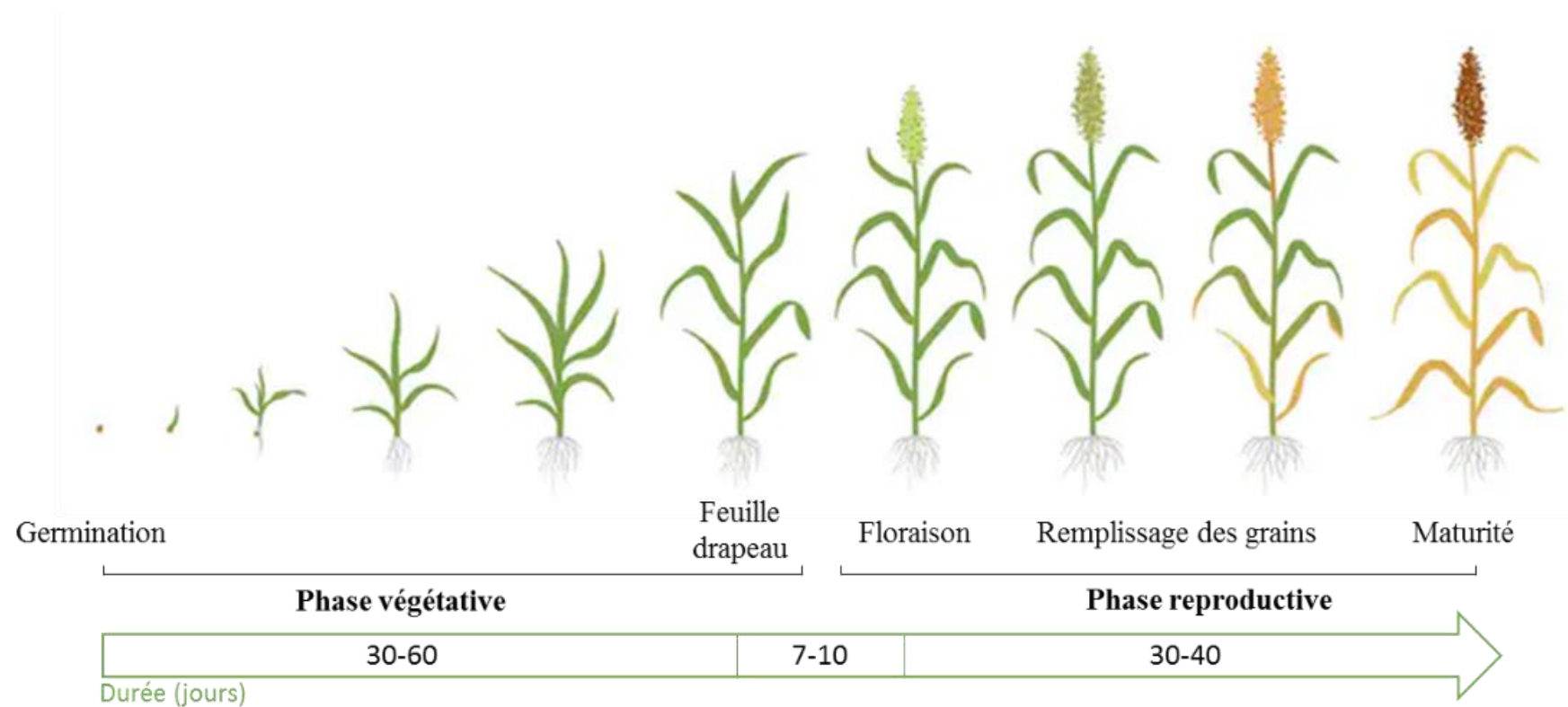


Figure I-3 : Stades de développement du sorgho de la germination à la maturité (stock illustration par Ilya Kalini)

Les axes portent les épillets (seul un ou deux sont fertiles) qui portent une fleur composée d'un ovaire à deux stigmates et trois étamines. L'ensemble des épillets portés par les branches de l'inflorescence forme la panicule. La forme et la compacité de la panicule sont diverses chez les sorghos cultivés et définissent par ailleurs les races botaniques (Figure I-2C). La taille du rachis et des ramifications détermine la compacité de la panicule. La floraison commence au sommet de la panicule et s'étend sur quelques jours jusqu'au bas de celle-ci. Le sorgho est une céréale monoïque et préférentiellement autogame avec en moyenne 6% de pollinisation croisées. Cependant, jusqu'à 30% d'allofécondation a été observée chez certaines variétés (Chantereau *et al.*, 1994).

Selon les variétés, la panicule porte 2000 à 4000 graines. La graine de sorgho contient un unique embryon et est protégée par un péricarpe plus ou moins épais et blanc, jaune ou rouge. Les glumes qui protègent les fleurs restent présentes sur les graines.

5. Cycle de développement

Le sorgho est une plante annuelle dont le cycle total s'étend de quelques mois à 6 ou 8 mois pour les variétés tardives (Figure I-3). Il comporte deux phases : la phase végétative qui dure de la germination à l'apparition de la feuille drapeau puis suivant 10 jours de développement de la feuille drapeau, la phase reproductive qui s'étale de la fécondation des fleurs jusqu'à la maturité des grains.

La phase végétative suit la germination, phase critique du développement dont les conditions de température idéales sont des moyennes journalières à 27°C (Chantereau *et al.*, 2013). Au stade juvénile, l'investissement de la plante est plus important dans ses racines que dans ses organes aériens. Chaque feuille apparaît avec un délai d'environ 4 jours selon la variété et les conditions climatiques. Ce délai est appelé phyllochrone et est habituellement mesuré en temps thermique (degrés jours) qui correspond à l'accumulation d'énergie thermique à partir de la température basale qui est la température minimum à laquelle l'espèce croît (chez le sorgho elle est de 11° en moyenne, (Kim *et al.*, 2010). A partir de 20 feuilles, et si la plante continue son développement, le phyllochrone peut doubler. L'arrêt de l'élongation des entrenœuds qui précède la transition du méristème végétatif vers un méristème reproducteur coïncide avec le début de l'accumulation des sucres solubles dans la tige. La phase végétative est la plus variable en termes de durée, allant d'un à plus de deux mois selon les génotypes. La phase reproductive qui suit est plus invariable et démarre à l'initiation florale. Cette phase est critique en terme de besoins en eaux, nutriments et lumière. Au total, 10 jours s'écoulent

entre le stade feuille drapeau et l'anthèse (floraison), la panicule apparaissant quelques jours après la feuille drapeau. Une fois la fécondation de la panicule complète débute la phase de remplissage des grains qui passent d'une couleur verte et d'une consistance laiteuse à une consistance pâteuse puis dure et une couleur blanche ou rougeâtre. Trente à 40 jours sont nécessaires pour arriver à la maturité physiologique depuis l'anthèse (Figure I-3).

6. Structure de la tige et organisation des parois

La biomasse sèche de la tige de sorgho représente en moyenne 30% de sa masse fraîche (Almodares et Hadi, 2009; Gutjahr *et al.*, 2013) et est principalement composée de la biomasse accumulée dans la tige sous forme de sucres solubles et de parois cellulaires. La tige des monocotylédones est un organe aérien rigide sans croissance secondaire qui a pour rôle de soutenir les parties aériennes, de stocker des éléments nutritifs, d'assurer leur transport entre les racines, les feuilles et les grains et de produire de nouveaux tissus. Elle est composée de tissus vasculaires qui transportent la sève et de fibres de structure qui assurent le soutien et la protection des vaisseaux conducteurs. L'épiderme, l'hypoderme et le parenchyme sont les tissus de soutien et le xylème et le phloème sont les vaisseaux assemblés en faisceaux vasculaires protégés d'une gaine fibreuse, orientés respectivement vers l'intérieur et l'extérieur de la tige (Figures I-4). La résistance mécanique de ces tissus est assurée par les parois secondaires qui les entourent (Smith *et al.*, 2017). La zone périphérique est 2 à 3 fois plus riche en parois végétales que la zone centrale et 3 à 4 fois plus riche en lignine, composé rigidifiant ces parois (Richie et McBee, 1991) (Figure I-4).

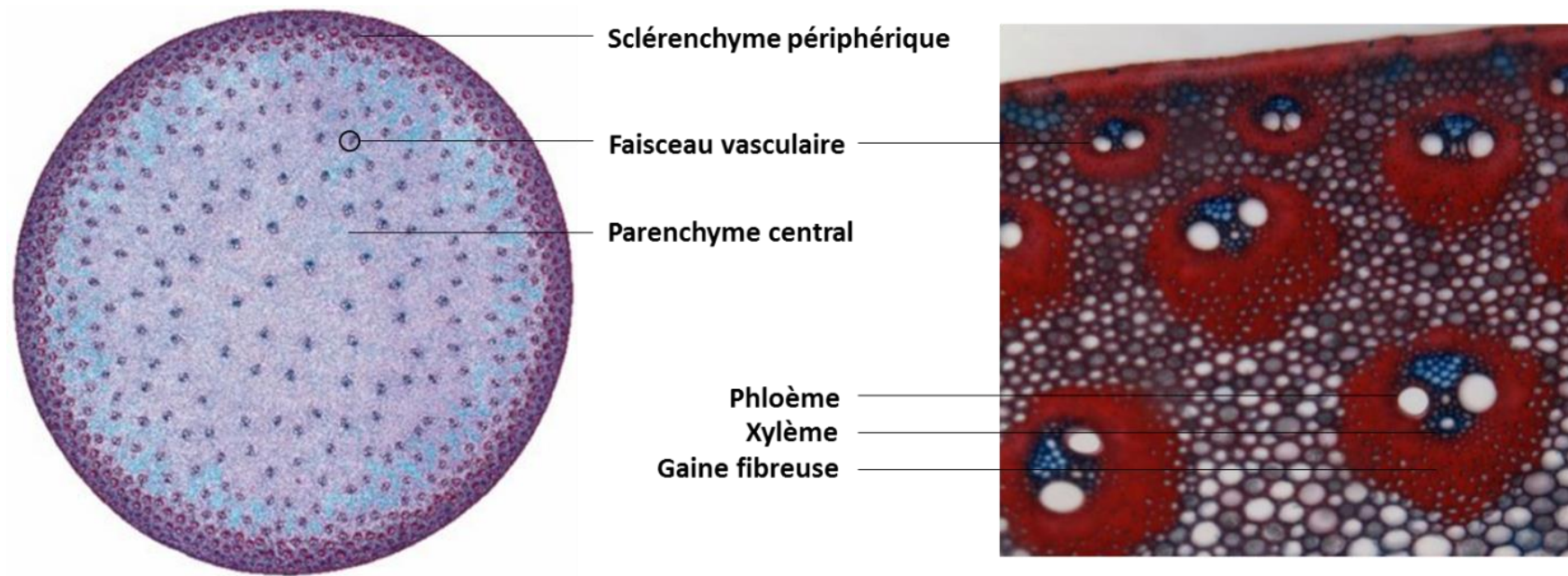


Figure I-4 : Coupe d'entre-nœud de sorgho colorée au Fasga : les lignines sont en rouge et les polysaccharides sont en bleu (D'après Perrier *et al.* 2017b et Jaffuel *et al.*, 2016)

Les parois secondaires des cellules de la tige constituent la majeure partie de la biomasse du sorgho. Elles se déposent autour des cellules après leur différenciation, leur élongation et le dépôt d'une paroi primaire. La paroi secondaire se dépose en trois couches (S1, S2 et S3) d'organisations différentes (Figure I-5) entre la membrane plasmique et la paroi primaire. La lamelle moyenne est le compartiment qui sépare deux cellules et qui est encadré par les parois primaires de ces cellules. La paroi primaire est une paroi souple qui permet la croissance et l'élongation de la cellule indifférenciée tandis que la paroi secondaire est rigide et imperméable et n'entoure que certains types de cellules après leur différenciation : les cellules conductrices de sèves et les cellules qui forment les tissus de soutien et de protection (Figure I-4). Les parois sont composées majoritairement de polysaccharides complexes (hémicelluloses, cellulose, pectines), de protéines et de polyphénols.

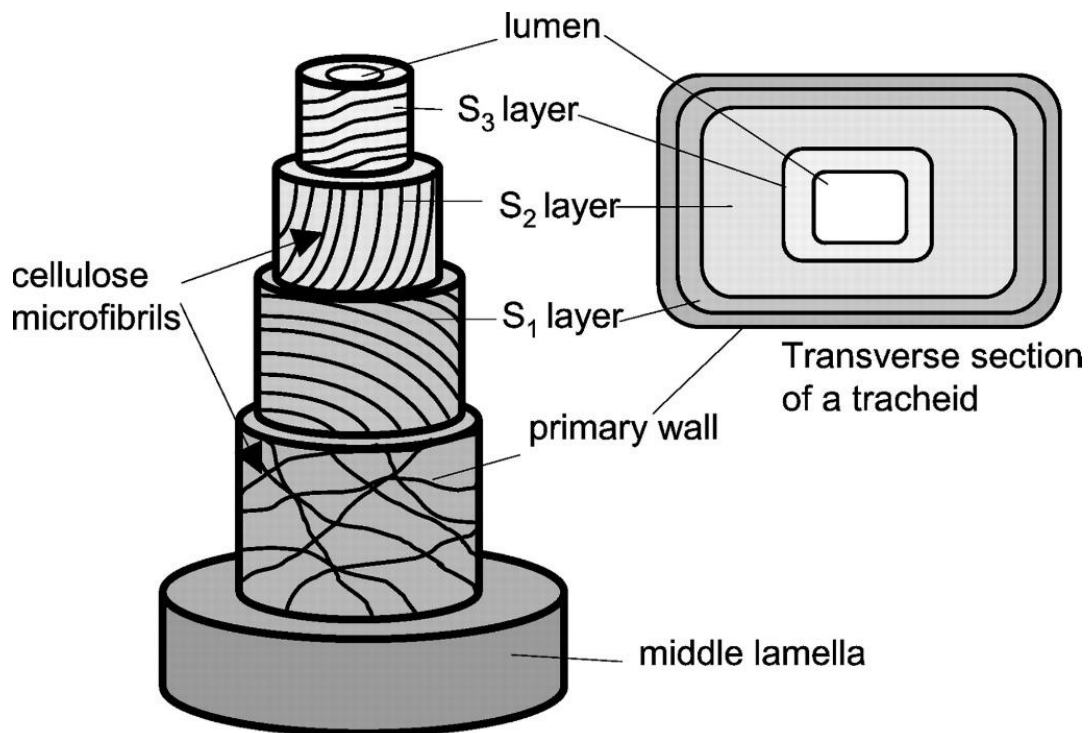


Figure I-5 : Organisation tridimensionnelle des parois et de la cellulose dans une cellule vasculaire du xylème. La paroi est divisée en différentes couches chacune ayant un arrangement spécifique des fibrilles de celluloses. Ces fibrilles peuvent être arrangées de façon irrégulière comme dans la paroi primaire ou à des angles spécifiques comme dans les couches de la paroi secondaire (D'après Plomion *et al.*, 2001)

Les différences de structure des éléments de la tige sont importantes entre dicotylédones et monocotylédones (Figure I-6). Chez les monocotylédones, les faisceaux vasculaires sont dispersés dans la tige sans l'arrangement en anneau qui existe chez les dicotylédones. Des faisceaux fibreux accompagnent les faisceaux vasculaires chez les monocotylédones alors qu'ils sont absents chez les dicotylédones. La moëlle, centrale chez les dicotylédones, est réduite chez les monocotylédones et c'est un parenchyme épais qui semble être responsable du transfert et du stockage des nutriments. Il n'y a pas de croissance secondaire chez les monocotylédones, ainsi aucun cambium ne s'intercale entre phloème et xylème chez les monocotylédones. D'autres différences existent entre les monocotylédones et les dicotylédones au-delà de la structure de la tige, elles sont indiquées au sein du Tableau I-1.

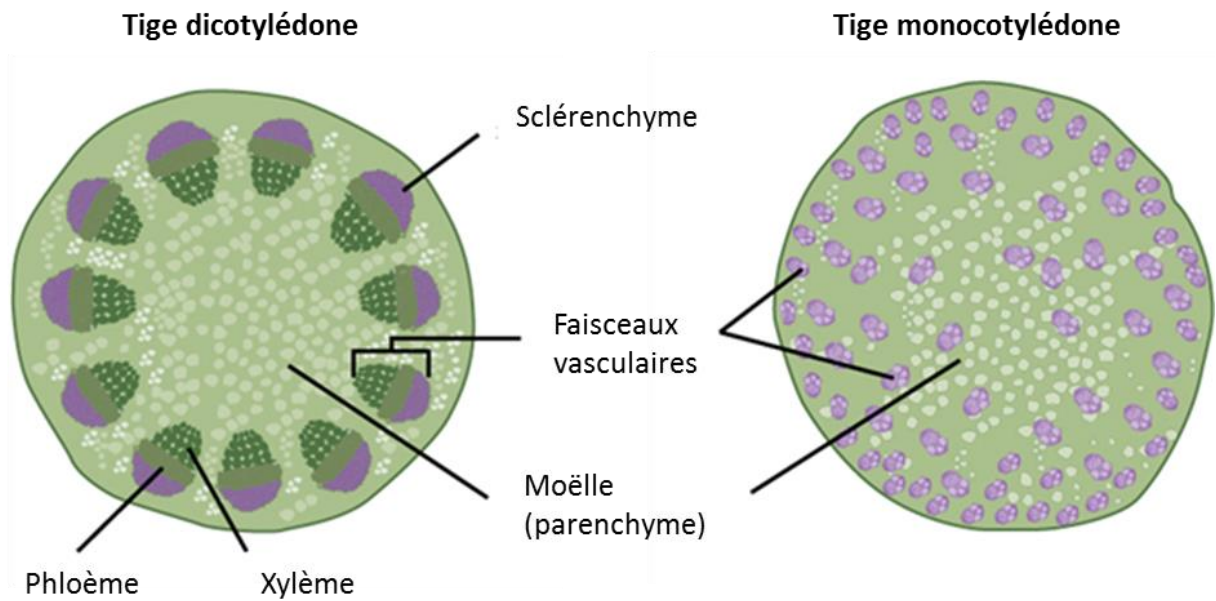


Figure I-6 : Différences de structure de la tige des dicotylédones et des monocotylédones (bio1520.biology.gatech.edu)

Tableau I-1 : Différences générales de morphologies entre monocotylédones et dicotylédones

Monocotylédones	Dicotylédones
Cotylédon unique	Deux cotylédons
Feuille longue à vasculature parallèle	Large feuille à vasculature centrale ramifiée
Parties florales en multiple de trois	Parties florales en multiple de 4 ou 5
Racines fasciculées (fibreuse et ramifié)	Racines pivotantes

7. Description des composés pariétaux

7.1. Cellulose

La cellulose est le composant principal des parois végétales et la forme principale de stockage du carbone sur Terre (Delmer et Haigler, 2002, Saxena et Brown, 2005). Elle est composée de polymères de β -1,4-glucanes construits à partir de molécules de glucose qui s'assemblent linéairement en β -(1 \rightarrow 4) avec une rotation de 180°C à chaque nouvelle molécule. Les dimères de glucose formés, le cellobiose, forment l'unité de base de la cellulose (Figure I-7). Cet assemblage et la polarisation qui en résulte lui confèrent des caractéristiques particulières. En effet, il permet de former un arrangement de longues chaînes qui peuvent être alignées côte à côte en créant des fibres d'une grande résistance mécanique. Dix-huit polymères de cellulose forment une micro-fibrille de cellulose. L'association de 6 micro-fibrilles de cellulose forme une macro-fibrille et un agencement de plusieurs macro-fibrilles forme une fibre de cellulose (Festucci-Buselli *et al.*, 2007) (Figure I-7). Le degré de polymérisation des chaînes de cellulose se situe autour de 2000 à 25000 glucoses (Brown, 1996). En conséquence de cet arrangement contrôlé, la résistance des fibres de cellulose est comparable à celle de l'acier (Eckardt, 2003). Selon la localisation dans le compartiment pariétal, la cellulose est présente sous forme de micro-fibrilles de 3 nm de diamètre ou de macro-fibrilles de 5 à 10 nm de diamètre (Delmer et Amor, 1995; Ding et Himmel, 2006; Jarvis, 2003).

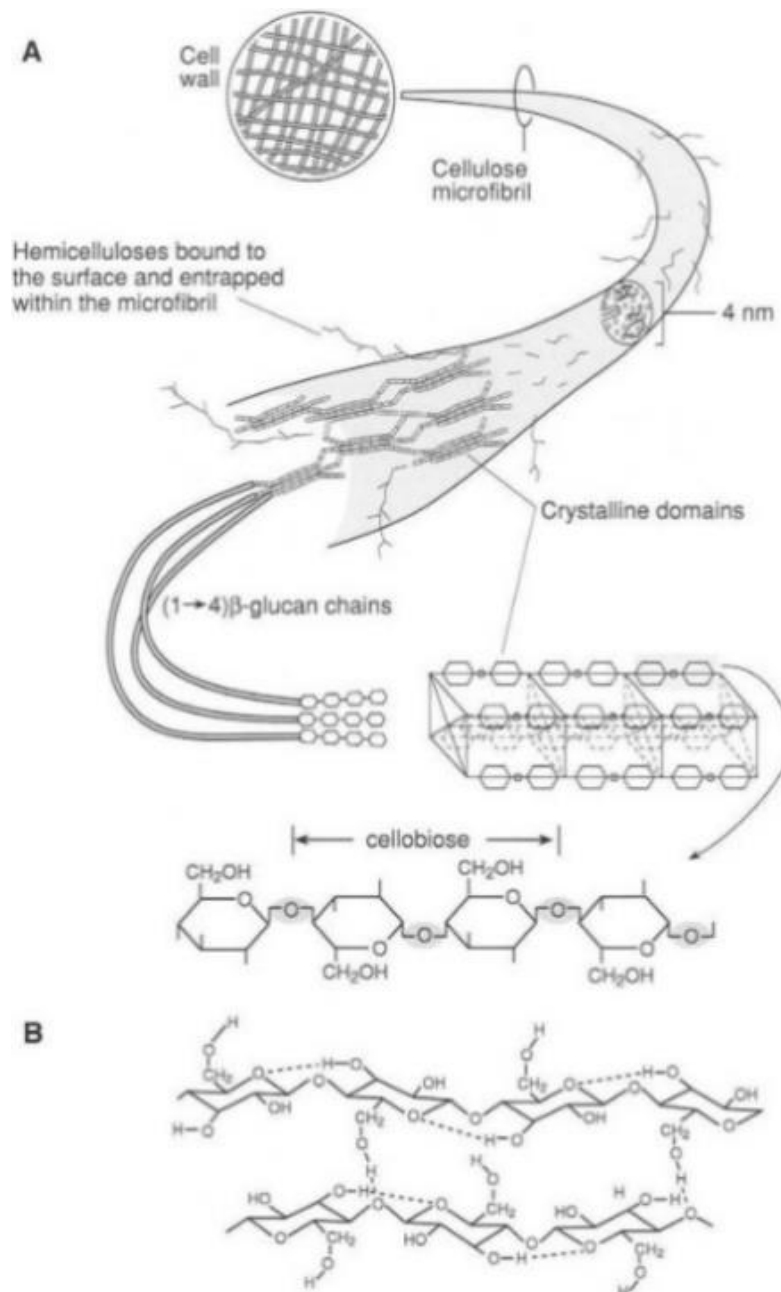


Figure I-7 : Structure d'une fibre de cellulose (A) Différents niveaux d'assemblage des polymères de glucose. La plupart des fibres de cellulose sont arrangées précisément en 3D et présentent des propriétés cristallines. (B) Deux chaînes de (1 → 4) β-glucane assemblées par liaisons hydrogènes (D'après Taiz et Zeiger, 1998)

7.2.Hémicelluloses

Les polysaccharides hémicellulosiques sont très hétérogènes et leur classification n'est pas clairement définie. Sont communément appelés hémicelluloses ; les xyloglucanes (XyG), les xylanes, les mannanes, les glucomannanes et les β -(1 \rightarrow 3,1 \rightarrow 4)-glucanes (Scheller et Ulvskov, 2010). Les hémicelluloses sont construites à partir d'une base d'oses liés équatorialement en β -(1 \rightarrow 4) qui ne s'assemblent pas en cristaux comme la cellulose et sont généralement acétylés (Figure I-8). Ils forment de longues chaînes qui s'associent à la cellulose notamment par des liens hydrogènes, hydrophobes et hydrophiles (Busse-Wicher *et al.*, 2014). Elles assurent la solidité de la structure complète de la paroi en retenant les microfibrilles de cellulose entre elles (Figure I-7). La classe, l'abondance et la ramification (méthylation, acétylation, substitutions de fucosyl, arabinosyl, galactosyl etc) des hémicelluloses varient beaucoup entre les espèces végétales, les types cellulaires et les stades de développement (Ebringerová, 2006; Pawar *et al.*, 2013). Leurs caractéristiques variées impactent différemment la saccharification de la biomasse lignocellulosique (Saha, 2003).

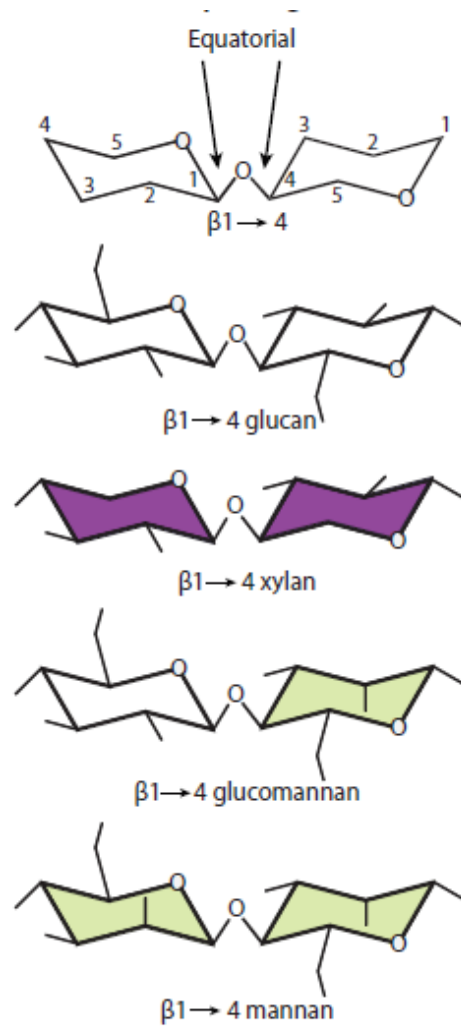


Figure I-8: Différentes unités de base des hémicelluloses (D'après Scheller et Ulvskov, 2010)

Les hémicelluloses peuvent être classifiées en 4 grands groupes (Figure I-9) :

- Les xyloglucanes (XyG) qui ont un squelette de résidus glucose sur lequel se greffent des résidus xylose, galactose et fucose. Les substitutions les plus courantes sont représentées en Figure I-8
- Les xylanes qui incluent :
 - les glucuronoxylanes (GX) qui ont un squelette de résidus xylose sur lequel se greffent des résidus d'acide glucuronique ou de son dérivé O-méthylé ;
 - les glucuronoarabinoxylanes (GAX) qui ont un squelette de résidus xylose sur lequel se greffent des résidus arabinose et acide glucuronique ;
 - les homoxylanes non substituées, les arabinoxylanes (AX) qui ont un squelette de résidus xylose sur lequel se greffent des résidus arabinose etc.
- Les mannanes qui ont un squelette de résidus mannose et les glucomannanes qui ont un squelette de résidus mannose et glucose, qui comprennent aussi :
 - les galactomannanes qui ont un squelette de mannose sur lequel se greffent des résidus galactose ;
 - les galactoglucomannanes qui ont un squelette de résidus mannose et glucose sur lequel se greffent des résidus galactose
- Les β -1,3;1,4-glucanes, encore appelés glucanes à liaisons mixtes (Mixed Linkage Glucan, MLG).

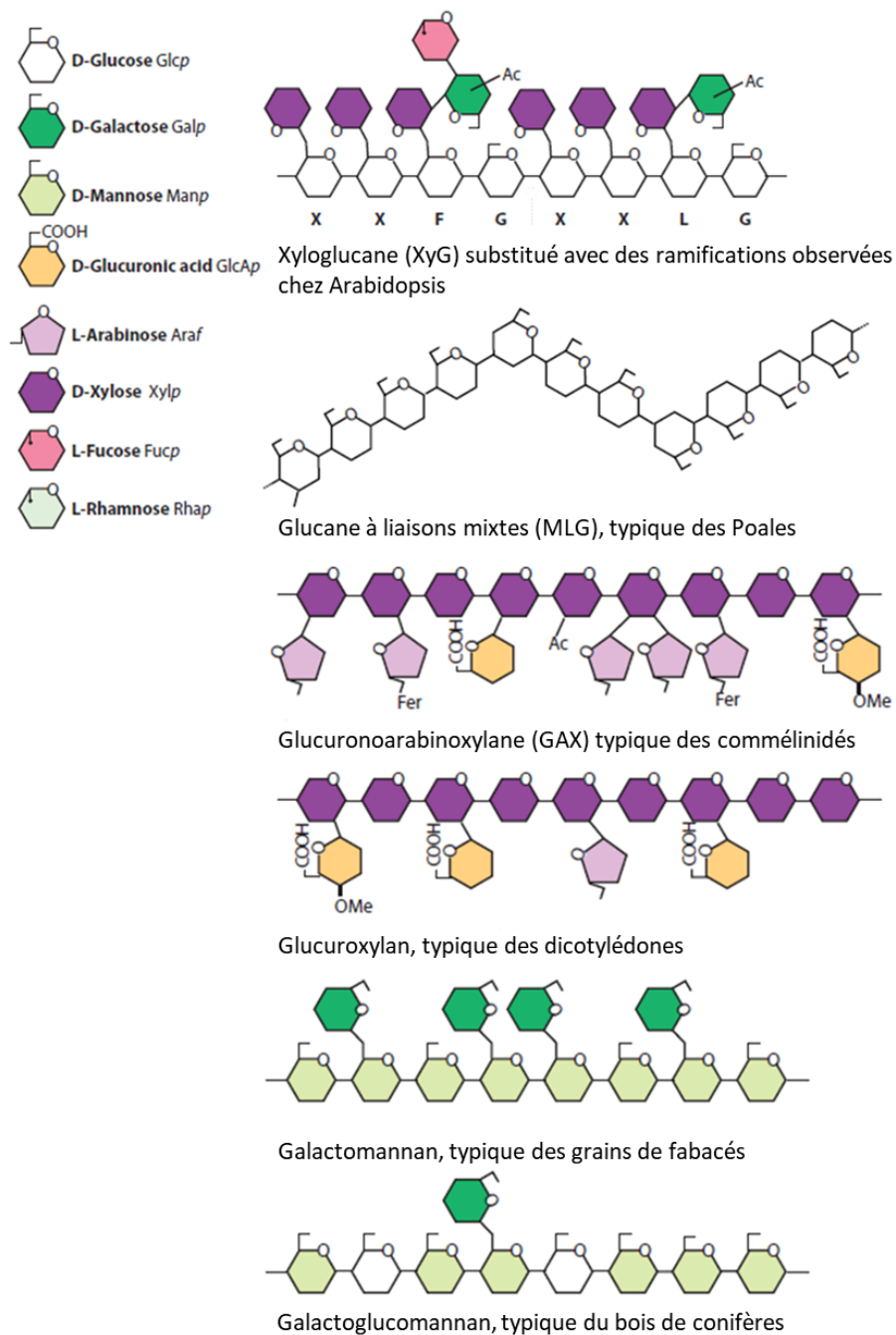


Figure I-9: Illustration des différents types d'hémicellulose présents dans les parois végétales. Les lettres présentes sous la molécule de xyloglucane sont les symboles utilisés pour les ramifications les plus communes. « Fer » représente l'estérification avec l'acide férulique qui est typique des xylanes des monocotylédones commelinidés. (D'après Scheller et Ulvskov, 2010)

7.3.Lignines

Les lignines sont des composés aromatiques issus de la polymérisation de trois monolignols qui en sont les unités principales : l'alcool p-coumarylique (H), l'alcool coniférylique (G) et l'alcool sinapylique (S)(Tableau I-2). Elles sont généralement absentes des parois primaires (Mellerowicz *et al.*, 2001). Les proportions des différentes unités H, G et S sont variables selon les tissus et les espèces (Lourenço et Pereira, 2018) et des unités de conformation différentes peuvent être incorporées à des niveaux plus faibles que les unités H, G et S selon l'espèce, le génotype et l'environnement (Ralph *et al.*, 2004).

La lignification commence dans les cellules vasculaires puis s'étend aux cellules fibreuses (Murakami, 1999). Elle s'initie au dépôt de la couche S1 des parois secondaires (Barros *et al.*, 2015; Terashima *et al.*, 1993) puis progresse à mesure que la couche S2 est déposée de façon concomitante avec les polysaccharides de structure. La lignification s'intensifie à la formation de la couche S3 et progresse vers le lumen jusqu'à ce que la lignification des parois soit complète (Mellerowicz *et al.*, 2001).

Les monolignols qui constituent la lignine se lient chimiquement de différentes manières entre eux et donnent à la lignine une hétérogénéité de structure (Boerjan *et al.*, 2003)(Figure I-10). D'autres composés phénoliques interviennent dans l'intégration des lignines aux parois secondaires. En effet, les acides hydroxycinnamiques (féruliques et *p*-coumariques) permettent aux lignines de s'éthérifier et s'estérifier aux hémicelluloses des parois végétales (Harris et Trethewey, 2010)(Figure I-11). La composition, les liaisons présentes et l'arrangement des lignines dans la matrice de polysaccharides sont déterminants dans la rigidité des parois secondaires et leur récalcitrance à la digestion (de Souza *et al.*, 2018; Molinari *et al.*, 2013; Tronchet *et al.*, 2010).

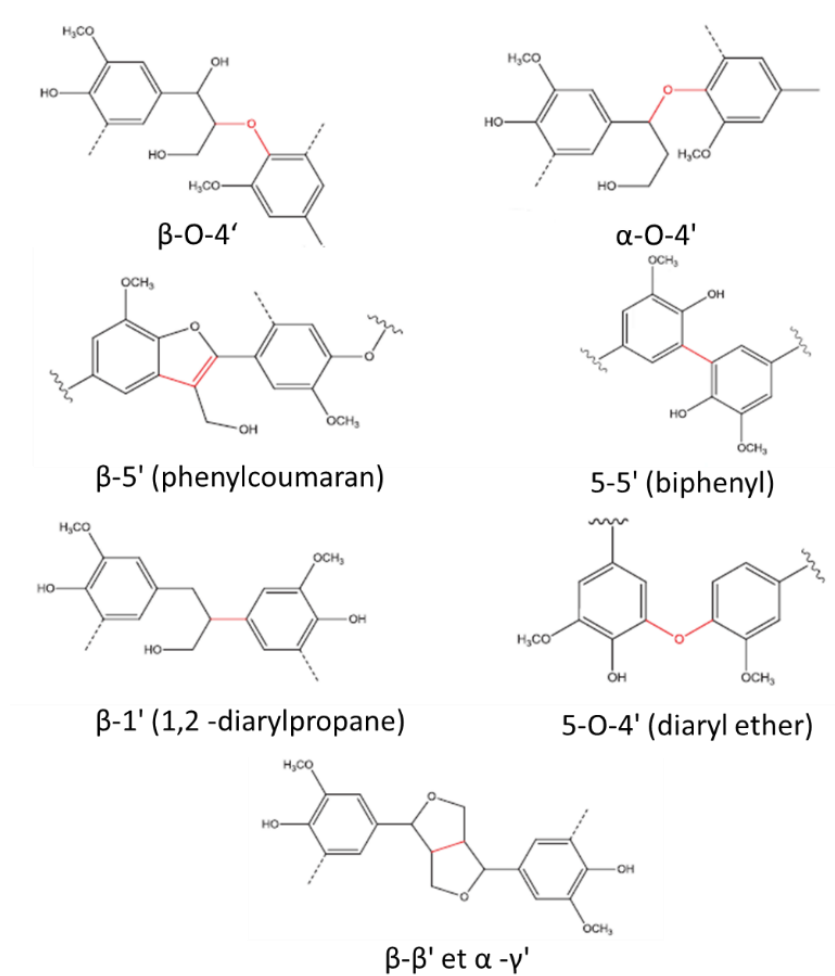


Figure I-10 : Différents types de liaisons présentes dans les polymères de lignine (D'après Lourenço et Pereira, 2018)

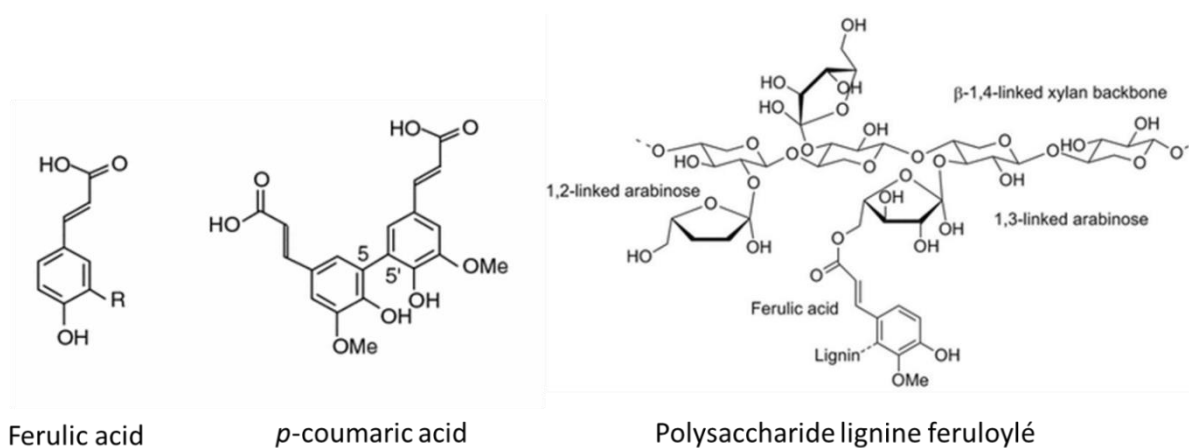
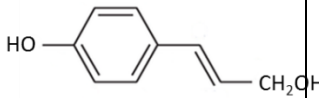
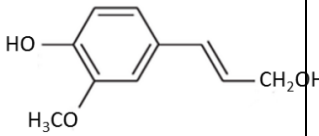
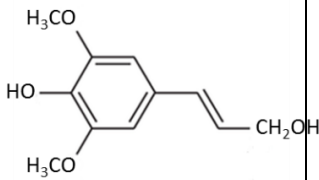


Figure I-11 : Composés phénoliques intervenant dans les liaisons de la lignine aux polysaccharides des parois (D'après Brandt *et al.*, 2013; Harris et Trethewey, 2010)

Tableau I-2 : Monolignols constitutifs des lignines (D'après Lourenço et Pereira, 2018)

Alcool	Unité	Symbole	Structure
p-Coumarylique	p-Hydroxyphényle	H	
Coniférylique	Gaïacyle	G	
Sinapylique	Syringyle	S	

7.4.Pectines

Les pectines sont un groupe de polysaccharides très hétérogènes. Les deux unités de base des pectines sont les homogalacturonanes (HG) et rhamnogalacturonanes I (RG I), les HG comptant pour 60% des pectines dans les parois (Ridley *et al.*, 2001)(Figure I-12). Les HG sont de longs homopolymères d'acide α -D-galacturonique dont les groupes carboxyl et hydroxyl sont méthyl-estérifiés et acétylés de façon variable. Les RGI sont des hétéropolymères disaccharidique d'acide galactopyranuronique et de rhamnose également méthyl-estérifiés (Ridley *et al.*, 2001). Des substitutions de ces unités centrales, en particulier HG, existent et sont composées de groupes méthyles ou acyles. Concernant RG I, ces chaînes sont exclusivement attachées aux rhamnoses. Ce sont des oligosaccharides de 4 types selon leur longueur ; arabinanes, galactanes ou arabinogalactanes (AG) de type I ou II (Caffall et Mohnen, 2009; Mohnen, 2008; Ridley *et al.*, 2001). Les ramifications de RG I sont spécifiques du tissu et du stade de développement ce qui suggère une spécialisation fonctionnelle (O'Neill *et al.*, 1990; Ochoa-Villarreal *et al.*, 2012).

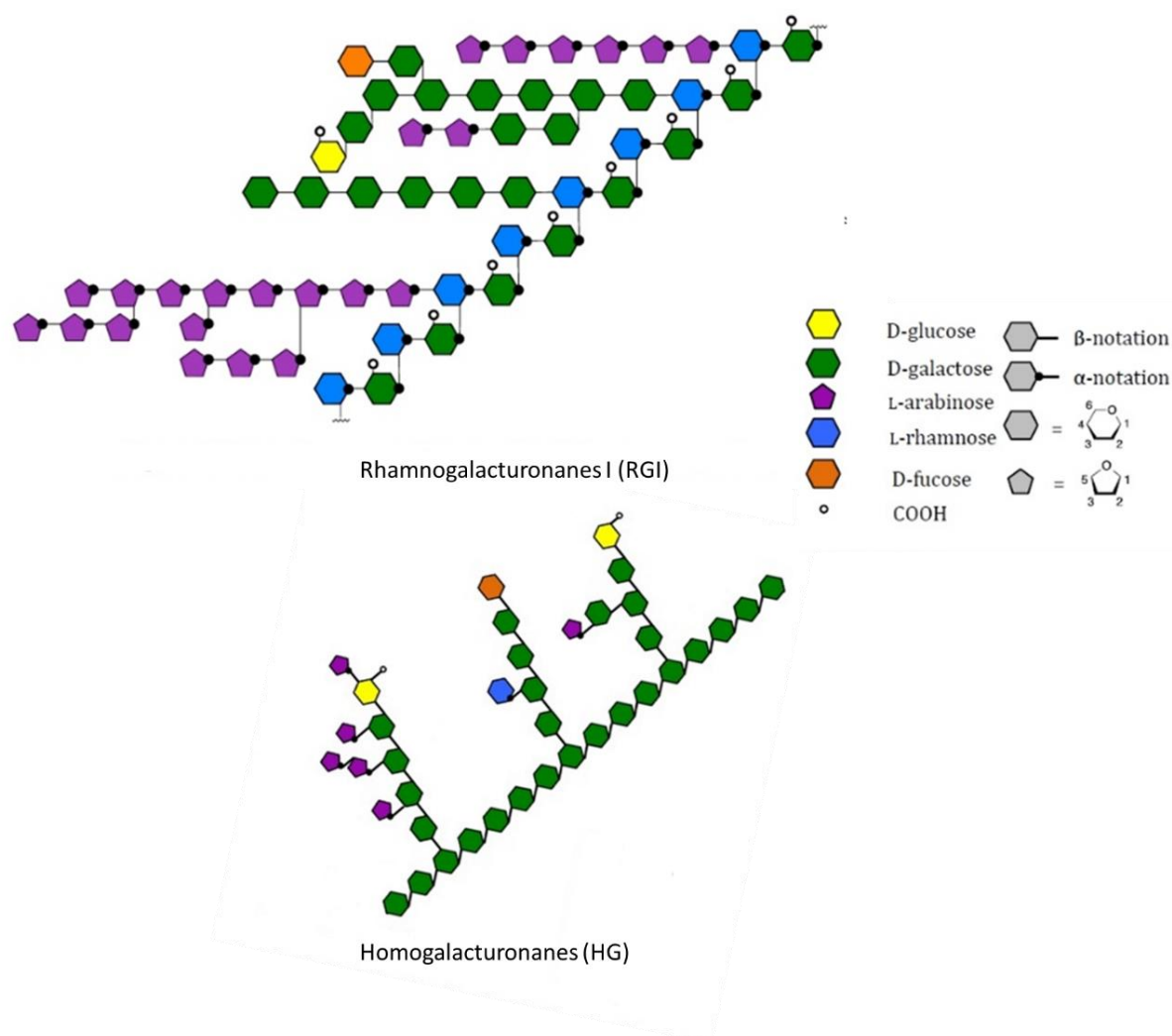


Figure I-12: Structure des deux unités de bases des pectines présentes dans les parois végétales (D'après Andersen, 2014)

7.5. Protéines pariétales

Les protéines incluses dans les parois végétales font parties de diverses familles. Les glycosides hydrolases (GH), les carbohydrates estérases/lyases et les expansines assurent le remodelage des polysaccharides et sont les plus abondantes dans les parois (Nguyen-Kim, 2015). Par exemple, les expansines semblent rompre des connexions entre les polysaccharides de la paroi primaire de cellules en expansion. Les oxydo-réductases dont les peroxydases et les blue copper binding proteins qui représentent 15% des protéomes pariétaux (Nersissian et Shipp, 2002, Cosgrove, 2018) interviennent dans le remaniement des réseaux pariétaux même si les mécanismes précis ne sont pas connus (Nguyen-Kim, 2015). Les protéases comme les Serine-carboxypeptidases (12% des protéomes pariétaux) pourraient être responsables de la maturation et la dégradation des protéines pariétales ainsi que de la signalisation peptidique (Van der Hoorn, 2008; Albenne *et al.*, 2009). De nombreuses autres protéines sont présentes dans les parois comme des protéines à domaine d'interaction qui interagissent avec des lectines ou d'autres protéines (LRR), des protéines de signalisation comme les protéines à arabinogalactanes (AGP), des fascilin-like AGP (FLA) qui pourraient contribuer aux propriétés mécaniques des parois végétales (Seifert et Blaukopf, 2010, Seifert et Roberts, 2007, Shiu et Bleecker, 2001). Les protéines liées au métabolisme des lipides pourraient participer à la liaison des lipides dans la cuticule et être impliquées dans l'extension de la paroi cellulaire en interagissant avec le réseau de polysaccharides (Samuels *et al.*, 2008; Yeats et Rose, 2013, Nieuwland *et al.*, 2005). Enfin, des protéines structurales sont potentiellement importantes pour la mise en place des parois mais peu représentées dans l'analyse des protéomes pariétaux à cause de leur liaison covalente à la paroi (Albenne *et al.*, 2014).

7.6. Flavonoïdes

Les flavonoïdes sont des métabolites secondaires appartenant à la classe des polyphénols. Ils sont la classe de métabolites secondaires les plus abondants chez les plantes, sont synthétisés dans tous les organes des plantes et sont responsables de la coloration et le goût des fruits, feuilles et graines (Corcoran *et al.*, 2012). Leur structure de base est proche car ils sont tous issus de la voie des shikimate et de la voie des phénylpropanoïdes (Figure I-13) (Tripoli *et al.*, 2007, Higashi et Saito, 2013). Les grandes classes de flavonoïdes sont les flavones, flavonols,

flavanones, dihydroflavonols, isoflavones, flavanols, anthocyanidines, chalcones et autones (Figure I-13).

S'ils sont synthétisés à l'intérieur des cellules (Jorgensen *et al.*, 2005), les flavonoïdes peuvent être transportés dans les parois (Agati *et al.*, 2012). Les flavonoïdes sont relâchés en tant que molécules de signalisation pour attirer ou repousser des prédateurs ou partenaires (Mierziak *et al.*, 2014). En particulier, ils sont incorporés dans les parois des cellules nécrotiques et leurs cellules adjacentes (Dixon 1995, Beckman *et al.*, 2000). Ils sont également impliqués dans l'inhibition des enzymes des pathogènes dont le rôle est la digestion des parois (Treutter, 2005) et dans le renforcement général des parois (Beckman *et al.*, 2000, Skadhauge *et al.*, 1997).

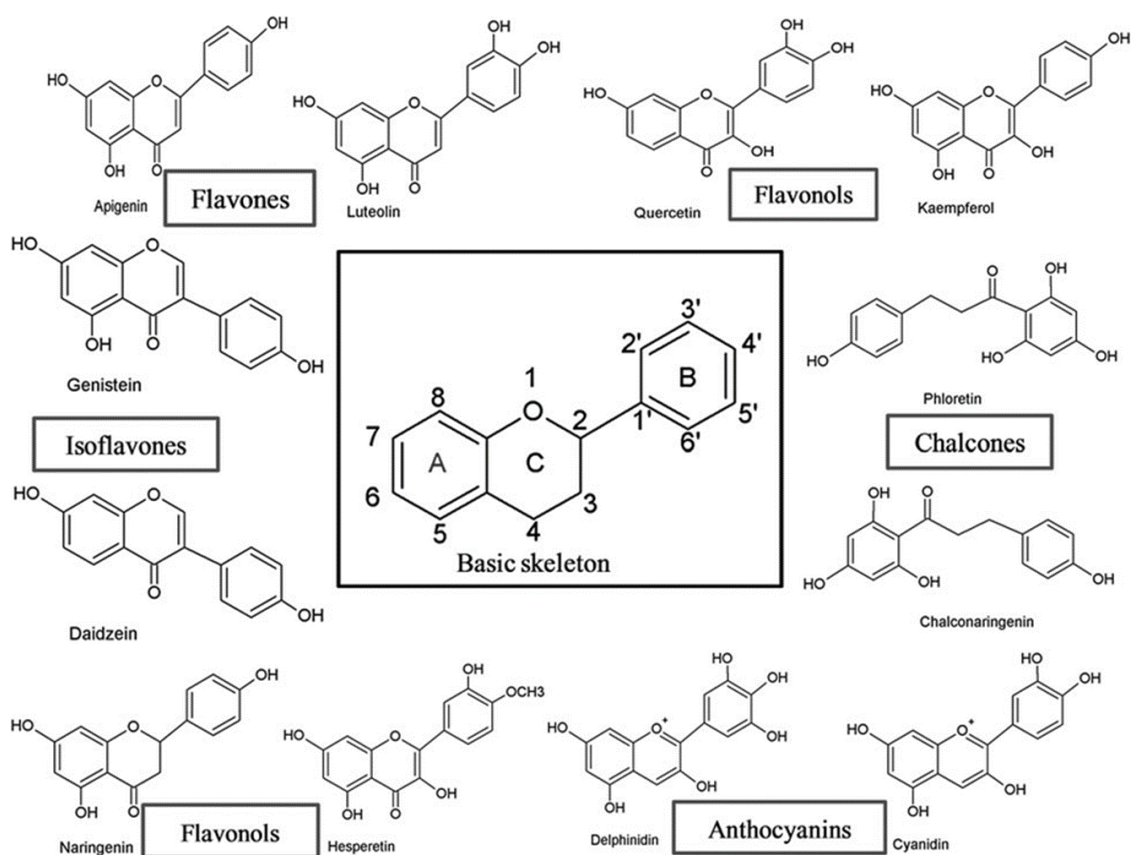


Figure I-13 : Squelette de base des flavonoïdes et grandes classes présentes chez les plantes (Panche *et al.*, 2016)

8. Composition générale des parois

La paroi primaire est composée à 90% de polysaccharides complexes et à 10% de composés protéiques (Figure I-14). La base de la paroi primaire est un réseau de fibrilles de cellulose inclus dans une matrice de polysaccharides non cellulosiques (hémicelluloses et pectines) (Figure I-15A). Les fibres de cellulose sont « cross-linkées » par les hémicelluloses (Cosgrove, 1997, Wolf *et al.*, 2012) et, même si leur interaction est moins claire avec les celluloses, les pectines assurent un rôle structurel important (Peaucelle *et al.*, 2012). La paroi primaire conserve des caractéristiques de flexibilité grâce à sa structure non lignifiée riche en pectines et en extensines. La paroi secondaire est une structure épaisse, au contraire de la paroi primaire, composée d'une matrice de cellulose et d'hémicellulose moins ramifiée et beaucoup plus ordonnée que dans les parois primaires (Figure I-15B). La paroi secondaire est une succession de couches de cellulose intégrées à un réseau d'hémicelluloses qui forment des liaisons covalentes avec les lignines (Albersheim *et al.*, 2011; Ebringerová *et al.*, 2005; Vogel, 2008). De surcroît, le dépôt de la paroi secondaire se fait en plusieurs couches orientées selon des angles différents, ce qui permet d'assurer un autre niveau de rigidité (Figure I-5). Ces couches (S1, S2 et S3) comportent des phases aux compositions et aux niveaux de cristallinité différents (Ruel *et al.*, 2012, Brett et Waldron, 1990). De plus, dans les parois secondaires, la cellulose s'assemble en macrofibrilles de 5 à 10nm de diamètre contrairement à la paroi primaire où les microfibrilles de 3 nm de diamètre prévalent (Delmer et Amor, 1995; Ding et Himmel, 2006; Jarvis, 2003, Festucci-Buselli *et al.*, 2007)(Figure I-7). Le niveau de polymérisation de la cellulose peut également doubler chez certaines espèces entre la paroi primaire et la paroi secondaire (Mellerowicz *et al.*, 2001).

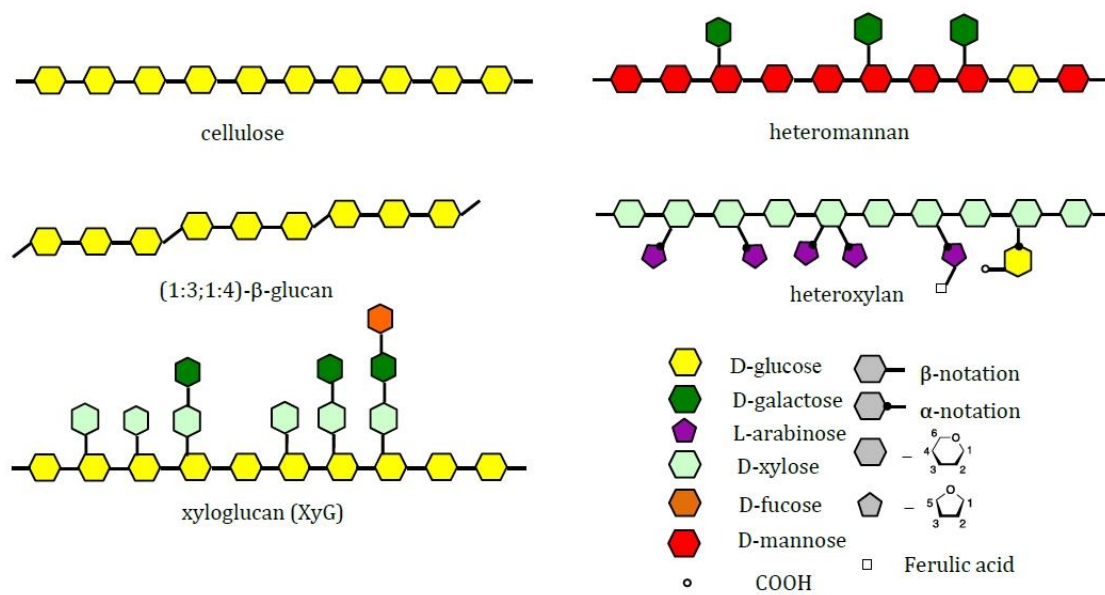


Figure I-14 : Structure des principaux polysaccharides des parois (D'après Andersen, 2014)

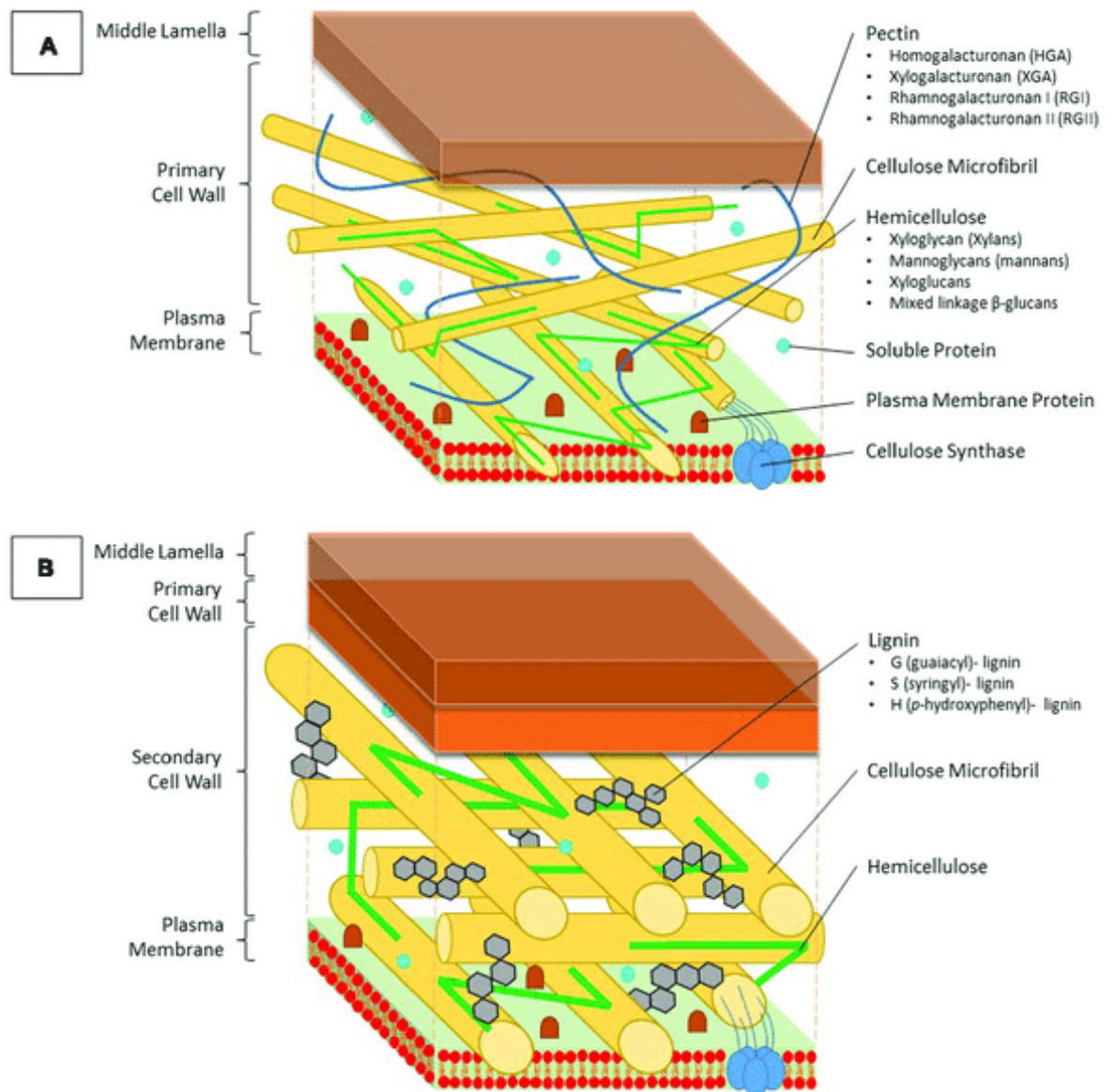


Figure I-15 : Structure et composition des parois primaires et secondaires chez les plantes. (A) La paroi primaire est construite entre la lamelle moyenne et la membrane plasmique et est composée de microfibrilles de cellulose dans une matrice d'hémicelluloses (dont les 4 types sont décrits), de pectines et de protéines pariétales. (B) La paroi secondaire est construite entre la paroi primaire et la membrane plasmique et est composée d'un arrangement très ordonné de macro-fibrilles de cellulose dans une matrice d'hémicelluloses et de lignines (D'après Loix *et al.*, 2017)

La structure et composition des parois végétales varie beaucoup entre monocotylédones et dicotylédones, entre paroi secondaire et parois primaires mais aussi entre espèces, tissus, stades de développement et selon l'état physiologique de la plante (Burton *et al.*, 2010).

Les composés présents dans les parois sont, pour la plupart, identiques entre dicotylédones et monocotylédones mais varient en termes de pourcentage de la composition totale des parois (Tableau I-4). En effet, les graminées présentent plus d'hémicelluloses que de cellulose et inversement pour les dicotylédones, les hémicelluloses MLG sont une spécificité des graminées et les mannanes sont très mineures chez les graminées alors qu'elles sont présentes chez les dicotylédones. Ces dernières sont des hémicelluloses typiques des parois secondaires des conifères.

Il existe deux types (1 et 2) de parois primaires qui varient pour la composition des hémicelluloses qui forment la matrice dans laquelle les fibres de cellulose sont incluses (Carpita 1996). Le type 1 est présent chez les dicotylédones et les monocotylédones non commélinées (les graminées font partie des commélinées) et est composé de pectines et de XyG en majorité avec une petite proportion de mannanes. Le type 2, présent chez les monocotylédones commélinées comme le sorgho, est composé de GAX, MLGs et d'une minorité de XyG (Doblin *et al.*, 2010, Burton *et al.*, 2010). Il contient moins de pectines que les parois primaires de type 1 et plus de phénylpropanoïdes (Iiyama *et al.* 1990) comme les composés hydroxycinnamiques (acides féruliques substitués sur les arabinoxylanes par exemple) (Nishitani et Nevins 1989, Ishii 1997).

Les parois secondaires des graminées sont enrichies en xylanes par rapport aux parois primaires tout comme les parois secondaires des dicotylédones. Celles-ci contiennent par contre peu de composés phénoliques alors qu'ils représentent entre 0.5 et 1.5% de la matière sèche des parois secondaires des graminées (Vogel, 2008, Scheller et Ulvskov, 2010).

Tableau I-4 : Composition caractéristique des parois primaires et secondaires des dicotylédones et des graminées en pourcentage de la matière sèche totale (D'après ^a Vogel, 2008 et ^b Scheller et Ulvskov, 2010, si non précisé les sources concordent ou l'une ne présente pas de données)

	Paroi primaire		Paroi secondaire	
	Graminées	Dicot	Graminées	Dicot
Cellulose	20-30	15-30	35-45	45-50
Hémicelluloses				
Xylanes	20-40	5	40-50	20-30
MLG	10-30 ^a /2-15 ^b	Absent	Mineur	Mineur
XyG	1-5	20-25	Mineur	Mineur
Mannanes	Mineur/2 ^b	5-10/3-5 ^b	Mineur/0-5 ^b	Mineur/0-7 ^b
Pectines	5	20-35	0.1	0.1
Protéines de structure	1	10	Mineur	Mineur
Phénols				
Acides férulique et p-coumarique	1-5	Mineur	0.5-1.5	Mineur
Lignines	Mineur	Mineur	20	7-10

9. Biosynthèse des principaux composants des parois secondaires

9.1. Biosynthèse de la cellulose

La biosynthèse de la cellulose est assurée par le complexe CSC (Cellulose Synthase), un complexe de protéines assemblées en rosette et composé de plusieurs sous unités composées elles-mêmes de plusieurs Cellulose-synthase A (CESA). Chaque sous-unité synthétise en parallèle des polymères de glucanes. Le nombre polymères synthétisés en parallèle, et donc le nombre d'enzymes impliquées, est débattu. Le modèle prévalent était de 6 sous-unités mais pourrait être de 18 sous-unités (Newman *et al.*, 2013; Nixon *et al.*, 2016, Brown et Saxena, 2000). Chaque sous-unité requiert la présence d'au moins 3 isoformes différentes de CESA (Doblin *et al.*, 2002; Ding et Himmel, 2006).

La cellulose est déposée sur la surface externe de chaque cellule et ne peut être synthétisée à l'intérieur du cytoplasme à cause de son insolubilité. Le complexe CSC est arrangé en forme de canal à l'interface de la membrane plasmique et intègre les UDP-glucoses dans les polymères de cellulose (Saxena et Brown, 2005 ; Carpita, 2011 ; McFarlane *et al.*, 2014 ; S. Li *et al.*, 2014). Plusieurs autres protéines sont impliquées dans la synthèse de la cellulose comme la P-sucrose synthase (P-SUSY) et l'endo-1,4- β -D-glucanase Korrigan (KOR) (Read et Bacic, 200, Delmer et Haigler, 2002; Molhoj *et al.*, 2002, Robert *et al.*, 2005, Vain *et al.*, 2014). P-SUSY pourrait être impliquée dans la conversion du sucrose en UDP-glucose et fructose, l'UDP étant relâché par l'assemblage du sucre dans la cellulose et disponible de nouveau pour la P-SUSY (Haigler *et al.*, 2001, Festucci-Buselli *et al.*, 2007). Les protéines COBRA, Chitinase-like et KOBITO sont également impliquées dans la synthèse de la cellulose sans que leurs rôles précis ne soient élucidés (Sadat Maleki *et al.*, 2016, Persson *et al.*, 2005, Zhang *et al.*, 2004, Zhong *et al.*, 2002, Schindelman *et al.*, 2001, Pagant *et al.*, 2002, Sanchez-Rodriguez *et al.*, 2012, Szyjanowicz *et al.* 2004, Vain *et al.*, 2014, Brown *et al.* 2005, Sato *et al.* 2010).

La synthèse de la cellulose dans la paroi primaire et la paroi secondaire est assurée par des groupes de CESA différents dont il existe 10 gènes au total dans le génome d'*Arabidopsis*. Parmi elles, CESA4, CESA7 et CESA8 ont été identifiées comme spécifiques de la biosynthèse de la cellulose dans les parois secondaires et CESA1, CESA3 et CESA6 dans les parois primaires chez *Arabidopsis* (Persson *et al.*, 2007, Tanaka *et al.*, 2003 Hernandez-Blanco *et al.*, 2007, Taylor *et al.*, 2003). Il existe des homologues de ces protéines dont le rôle

a également été démontré chez le riz, le peuplier et brachypodium (Tanaka *et al.* 2003, Joshi *et al.* 2011, Handakumbura *et al.* 2013).

9.2.Biosynthèse des hémicelluloses

Contrairement aux celluloses, les hémicelluloses sont synthétisées dans l'appareil de Golgi puis transportées vers les parois grâce à des vésicules membranaires (Sinclair *et al.*, 2018). Les différentes familles d'hémicelluloses sont synthétisées par des Glycosyltransférases (GT) appartenant aux superfamilles des cellulose synthases et des cellulose synthases-like dont font également partie les CESA responsables de la synthèse de la cellulose. Les squelettes des xyloglucanes (XyG) sont synthétisées par les CSLC (Cellulose-synthase-like C) puis subissent d'autres modifications avant d'être transportées à la paroi (hydrolyse, acétylation, méthylation, Gille *et al.*, 2012, Scheller et Ulvskov, 2010, Perrin *et al.*, 1999). La biosynthèse des mannanes est assurée par les CSLA (Cellulose-synthase-like A) et peut-être les CSLD (Cellulose-synthase-like D) (Dhugga *et al.*, 2004, Liepman *et al.*, 2005). La biosynthèse des MLG implique des protéines appartenant à la famille des cellulose-synthases qui sont absentes chez *Arabidopsis* et le peuplier mais présentes chez le riz et brachypodium (Scheller et Ulvskov, 2010). Ce sont les CSLF (Cellulose-synthase-like F) et les CSLH (Cellulose-synthase-like H) (Burton *et al.*, 2006, Doblin *et al.*, 2009). Les xylanes ne sont pas synthétisés par des GT de type I comme les autres hémicelluloses mais par des GT de type II des familles GT43, GT8, GT47 (IRREGULAR XYLEM8 (IRX8), IRX14, IRX10) (Rennie et Scheller, 2014, Scheller et Ulvskov, 2010). La biosynthèse de ces deux dernières classes d'hémicelluloses (MLG et xylanes) reste mal comprise.

9.3.Biosynthèse des lignines

La biosynthèse des lignines est généralement séparée en deux : la voie des phénylpropanoïdes qui part de la phénylalanine au feruloyl-CoA et la voie spécifique des monolignols du feruloyl-CoA aux monolignols (Figure I-16). Les acteurs de la biosynthèse des lignines sont relativement bien connus, en particulier chez l'espèce modèle *Arabidopsis thaliana* où la plupart des gènes a été clonée et étudiée (Raes *et al.*, 2003). Les gènes impliqués ont également été identifiés en partie chez les monocotylédones. Une dizaine d'étapes enzymatiques intervient dans l'ensemble des deux voies, phénylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate:CoA ligase (4CL), hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT), p-coumarate 3-hydroxylase (C3H), caffeoyl shikimate esterase (CSE) caffeoyl-CoA O-methyltransferase (CCoAOMT),

cinnamoyl-CoA reductase (CCR), ferulate 5-hydroxylase (F5H), caffeic acid O-methyltransferase (COMT), et cinnamyl alcohol dehydrogenase (CAD) (Boerjan *et al.*, 2003; Bonawitz et Chapple, 2010, Vanholme *et al.*, 2013). Après leur synthèse, les monolignols (p-coumaryl alcohol (unité p-hydroxyphenyls (H)), coniferyl alcohol (unité guaiacyl (G)), et sinapyl alcohol (unité syringyl (S))) sont secrétés à travers la membrane plasmique vers la paroi végétale (Wang *et al.*, 2013) où ils sont oxydés par des radicaux puis combinés (Vanholme *et al.*, 2012). Une voie parallèle dans le cytosol a été mise en évidence récemment. Elle permettrait de constituer une voie parallèle à la voie générale en amont de la 4CL chez les dicotylédones et est une étape indispensable de la voie de biosynthèse des monocotylédones (Barros *et al.*, 2019, Vanholme *et al.*, 2013).

Ce processus combinatoire de centaines de monomères par molécule de lignine est original car il ne serait pas catalysé par des enzymes. La régulation de ce processus par des laccases et des peroxydases est débattue mais serait largement chimique (Davin et Lewis, 2000, Boerjan *et al.*, 2003, Berthet *et al.*, 2011, Fraser et Chapple, 2011) malgré des évidences fonctionnelles du rôle des enzymes qui se sont récemment accumulées (Tobimatsu et Schuetz, 2019).

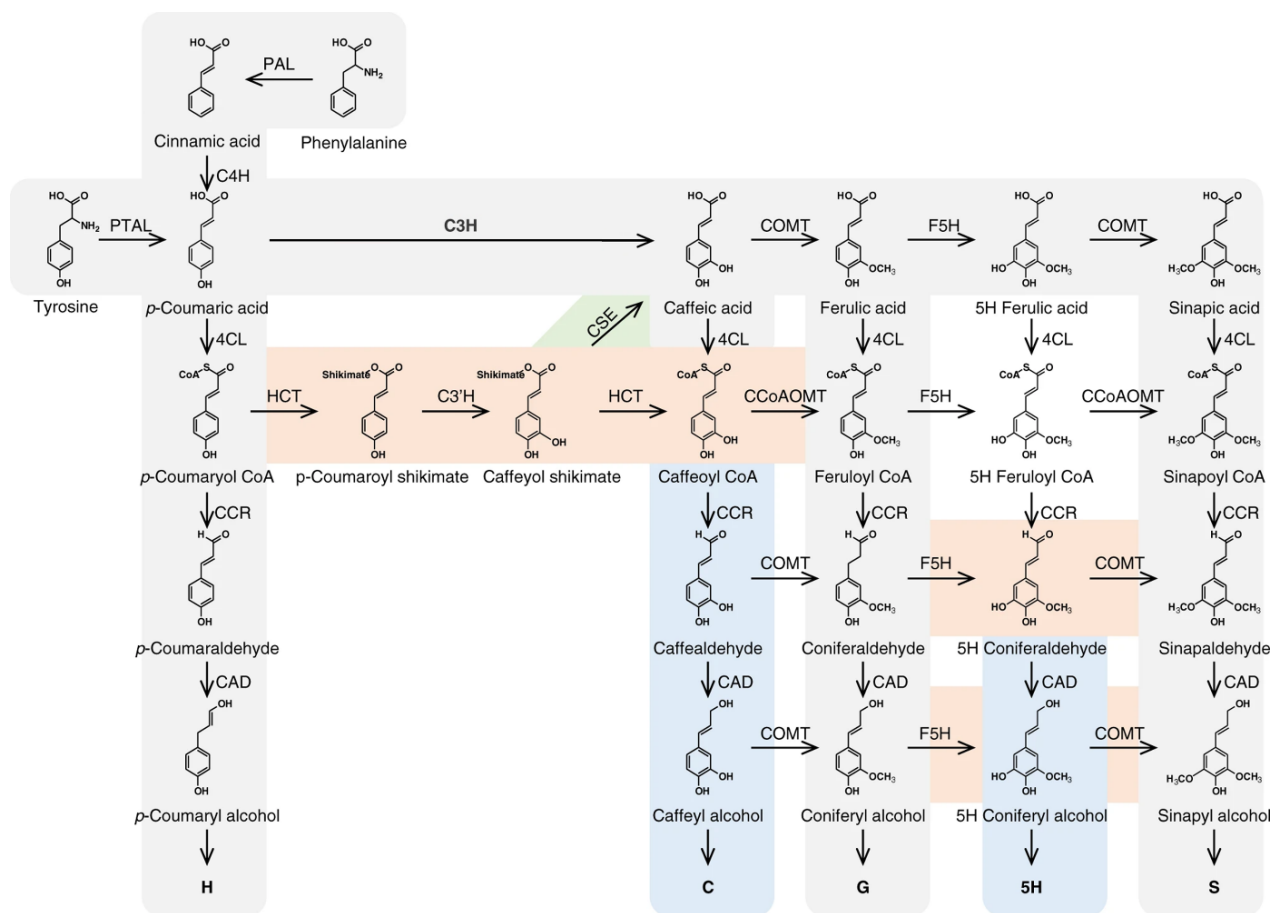


Figure I-16 : Voie de la biosynthèse des lignines. Les couleurs indiquent les étapes de la compréhension de la voie de biosynthèse. La voie telle que décrite dans les années 1980 (gris) suivi de la découverte dans les années 1990s de la dérivation par le shikimate (marron) puis celle des monomères de lignines C et 5H (bleu) et enfin la plus récente découverte de l'enzyme CSE (vert) (D'après Barros *et al.*, 2019)

10. Variation de la composition des parois en réponse aux « stress »

Les parois végétales évoluent au cours de la vie de la plante et peuvent être remodelées en situation de stress. En effet, la capacité des plantes à mobiliser des acteurs de la modification de leurs parois permet des changements structuraux compatibles avec l'adaptation à différents stress (Tenhaken, 2015).

10.1. Stress abiotiques

La sécheresse, le stress osmotique et l'exposition à un fort taux de sel entraînent la plupart du temps l'arrêt de la croissance des plantes (Skirycz et Inze, 2010). En effet, la turgescence cellulaire permise par l'absorption d'eau est un des principaux moteurs de la croissance (Schopfer, 2006). La croissance des parties aériennes s'interrompt avant celle du système racinaire permettant à la plante d'investir ses ressources restantes dans l'exploration du sol à la recherche de l'eau résiduelle (Sasidharan *et al.*, 2011). Chez le maïs, la croissance est restreinte à l'extrémité de l'apex racinaire où les derniers 2 à 6 mm de cellules ont une extensibilité des parois supérieure aux régions dont la croissance est stoppée. L'arrêt de croissance est accompagné de l'augmentation de l'expression de gènes impliqués dans la biosynthèse des lignines (Cinnamoyl-CoA) et une augmentation du dépôt de la lignine (Fan *et al.*, 2006). La susceptibilité à l'action des extensines est également inférieure dans ces régions, indiquant qu'un changement de structure des parois a bien eu lieu (Wu *et al.*, 1996).

Si les endoglucanases, XTH (xyloglucan endotransglucosylases/hydrolases) et extensines, responsables du remodelage polysaccharides, sont induites par le stress osmotique dans certaines régions des racines (Zhu *et al.*, 2007), les peroxydases le sont également (Maia *et al.*, 2013, Ranjan *et al.*, 2012, Pechanova *et al.*, 2010). Le rôle décrit de ces dernières est le cross-linking des hémicelluloses avec des composés phénoliques comme l'estérification des acides féruliques aux arabinoxylanes des graminées (Lindsay et Fry, 2007, Burr et Fry, 2009). Elles sont donc impliquées dans la rigidification des parois en réponse au stress hydrique. Elles pourraient également permettre l'élimination du peroxyde d'hydrogène, un radical libre (ROS) qui endommage les cellules et est associé à la réponse aux stress (Miller *et al.*, 2010, Tenhaken, 2015). Le renforcement de la paroi en plus de l'élimination des ROS pourrait être bénéfique pour la tolérance des structures cellulaires aux changements de turgescence induits par le stress osmotique (Kieffer *et al.*, 2000, Wakabayashi *et al.*, 2012). Cependant, il est aussi suggéré que la consommation des substrats des peroxydases mène à l'accumulation de ROS ce qui favorise l'apparition de radicaux OH responsables, eux, de la rupture des polymères

des parois et leur affaiblissement mécanique (Fry, 1998). De plus, les XTH et les expansines sont associées, tout comme dans la racine, à des conditions de stress osmotiques dans la tige (Rose *et al.*, 2002, Todaka *et al.*, 2012). Ces protéines promeuvent le relâchement des parois en conditions normales et permettraient la récupération post-stress hydrique (Tenhaken, 2015).

En apparence contradictoire, l'assouplissement de la paroi s'intègre dans un modèle où, après une période d'arrêt de croissance et de rigidification des parois pour faire face à la baisse de pression osmotique, les cellules croissent de nouveau grâce à l'affaiblissement des parois qui permet la croissance par turgescence (Tenhaken, 2015).

Les pectines sont également impactées par la sécheresse. La formation de gels hydratés par l'augmentation de leurs ramifications serait associée à la réduction des dommages cellulaires induits par le stress (Leucci *et al.*, 2008, An *et al.*, 2014). Elles sont également modifiées par le froid qui est associé à une augmentation du contenu en pectines et en activité des Pectin Methyl Esterases (PME) (Solecka *et al.*, 2008). L'activité des PME correspond à l'augmentation du niveau de des-estérification des pectines ce qui permet leur cross-linking par des ions calcium entraînant la formation d'un gel de pectines rigide (Sasidharan *et al.*, 2011). Tout comme pour le stress hydrique, la rigidification des parois fait partie des mécanismes de l'acclimatation au froid. En effet, le froid entraîne la formation de cristaux de glace dans l'apoplaste des cellules chez les espèces sensibles (dont font partie les espèces tropicales) tandis que les espèces résistantes restreignent le gel aux compartiments extra-cellulaires (Xin et Browse, 2000, Yamada *et al.*, 2002). En plus du stress induit par la déshydratation causée par le déséquilibre osmotique, les cristaux endommagent les matrices extracellulaires (Rajashekar et Burket, 1996, Smallwood et Bowles, 2002). Les parois rigides préviennent sa déformation et l'effondrement de la cellule à cause de la déshydratation. Les températures froides induisent une augmentation de l'épaisseur des parois et de leur rigidité à laquelle contribuent les pectines (Stefanowska *et al.*, 1999, Rajashekar et Burke, 1996, Rajashekar et Lafta, 1996). La croissance et l'expression des expansines et XTH, enzymes favorisant l'extensibilité des parois en condition de déficit thermique, sont également réduites.

La composition en lignines des parois ne semble pas modifiée par le froid (Solecka *et al.* 1999). Cependant, l'activité des enzymes de la voie de synthèse des monolignols (C3H, PAL) est augmentée et les niveaux d'acide p-coumarique et acide férulique sous forme soluble peuvent être également affectés (Janas *et al.*, 2000, El Kayal *et al.* 2006). Ces acides estérifiés pourraient être des substrats pour la polymérisation de la lignine mais pourraient aussi

permettre l'élimination des ROS en tant que substrats de peroxydases (Moura *et al.*, 2010). En effet, l'activité des peroxydases est augmentée en réponse au froid (Janda *et al.* 2007, Jain *et al.* 2007). D'une manière générale, les lignines induites par les stress comme le froid et la sécheresse se retrouvent dans les parois primaires des tissus normalement non lignifiés. Les ratios des différentes sous-unités sont également modifiés (moins de S et plus de G et H) tout comme la condensation des liaisons (Barros *et al.*, 2015).

Comme beaucoup des modifications des parois, celles induites par le froid sont potentiellement irréversibles. La rigidification des parois participe donc probablement à l'effet d'habituation au froid observé chez de nombreuses espèces qui survivent mieux à des épisodes de gels après des expositions à des températures basses (Chinnusamy *et al.* 2007).

La tolérance à la submersion peut impliquer l'acidification de l'apoplaste (Vreeburg *et al.*, 2005, Cosgrove, 2000) qui permet l'action d'enzymes modificatrices des parois et la mobilisation des expansines dont l'action sur l'extensibilité des parois entraîne une rapide elongation des entre-nœuds et l'émergence de la plante à l'air libre (Voeselek *et al.*, 2006, Kende *et al.*, 1998, Cho et Kende, 1997). Certaines réponses au stress de submersion entraînent au contraire un arrêt de la croissance via la régulation négative des gènes impliqués dans les processus énergivores comme la mise en place des parois, incluant donc les expansines (Fukao *et al.*, 2006). Cette modulation des parois n'est pas possible indéfiniment car après un épisode de submersion la susceptibilité des parois à l'action des expansines est réduite. Cela pourrait s'expliquer par un changement dans l'accessibilité des polysaccharides suite au remodelage de la paroi. En effet, une augmentation de la proportion de certains oses, polymères non cellulotiques et acides phénoliques est observée (Sasidharan *et al.*, 2011).

10.2. Stress biotiques

Les parois secondaires forment une barrière physique à l'entrée des organismes pathogènes et contribuent largement à l'immunité de la plante (Cantu *et al.*, 2008, Underwood, 2012). Des modifications dans le dépôt des parois secondaires ou dans la biosynthèse et la mise en place des polysaccharides de structure ont un impact important dans la susceptibilité des plantes aux attaques pathogènes (Miedes *et al.*, 2014). Les pathogènes utilisent souvent leur propre arsenal de protéines modificatrices des parois pour les dégrader et accéder aux tissus végétaux (Esquerre-Tugaye *et al.*, 2000). La lignine a un rôle important dans la constitution de cette barrière (Buendgen *et al.*, 1990) et un stress biotique induit la synthèse

rapide et le dépôt de lignines ou de composés proches dans les parois (Tronchet *et al.*, 2010, Sattler et Funnell-Harris, 2013). De façon attendue, les gènes de la voie des phénylpropanoïdes sont effectivement induits par les stress biotiques (Bhuiyan *et al.*, 2009). La lignine n'est cependant pas le seul produit de cette voie, de nombreux composés phénoliques comme les stilbènes, coumarines et flavonoides en dépendent (Dixon *et al.*, 2002, Naoumkina *et al.*, 2010). La plupart d'entre eux sont considérés comme des phytoalexines, des composés antimicrobiens impliqués dans les mécanismes de défense (Daayf *et al.*, 2012, König *et al.*, 2014). Les plantes synthétisent également des enzymes de dégradation à destination de leurs pathogènes comme des chitinases et des glucanases (Veronese *et al.*, 2003, Thomma *et al.*, 2002). Leur accumulation extra-cellulaire (Hamid *et al.*, 2013) implique une interaction avec les parois végétales.

11.Déterminisme génétique et évaluation de la qualité de la biomasse

La qualité de la biomasse dépend principalement de deux composantes : sa teneur en sucres solubles et la composition des parois cellulaires en polysaccharides de structure, en lignines et en autres composés comme les acides hydroxycinnamique. L'organe qui contribue le plus à la biomasse végétative est la tige, la qualité et le rendement en biomasse de la tige de sorgho sont donc primordiaux. Les sucres solubles sont synthétisés grâce à la photosynthèse et stockés dans le parenchyme de la tige. Ils se trouvent sous la forme de sucres simples (glucose, fructose) ainsi que sous forme de saccharose, un sucre de réserve. Les sucres de structures sont eux localisés dans les parois cellulaires sous formes de longs polysaccharides (cellulose, hémicelluloses, etc.). Selon les variétés et l'environnement, le pourcentage de matière sèche représenté par les sucres solubles varie d'un facteur 10 (Perrier, 2017a). Les différents idéotypes de sorgho varient également énormément dans leur composition en composants pariétaux (Mcbee et Miller, 1982; Monk *et al.*, 1984; Slewinski, 2012; Wang *et al.*, 2009). Globalement cependant, la cellulose représente 30 à 45% de la masse sèche des tiges de sorgho, les hémicelluloses 15 à 30% et les lignines de 5 à 20% (Kim et Day, 2011 ; Nasidi *et al.*, 2015 ; Vogel *et al.*, 2008, Trouche *et al.*, 2014)

La biomasse constituée par les parois secondaires est qualifiée de ligno-cellulosique et représente la plus grande part de la biomasse végétative, c'est une des principales ressources de l'agriculture. L'optimisation de la transformation de la tige est intrinsèquement liée à la qualité de sa biomasse et les propriétés optimales de celle-ci dépendent des produits finaux ciblés. En

effet, les fourrages et ensilages destinés à l'alimentation animale sont bien valorisés lorsque la biomasse est riche en énergie et surtout digestible. L'accessibilité des sucres est également recherchée dans les filières de conversion en bio-énergie (Lin et Tanaka, 2006). Au contraire, les sorghos destinés au renforcement des bio-matériaux profitent d'une biomasse dont la structure est difficile à altérer (Vo *et al.*, 2017). La modulation contrôlée de la qualité de la biomasse ligno-cellulosique est donc cruciale pour l'adaptation du sorgho à ces usages. A fortiori en conservant en parallèle les qualités agronomiques du sorgho comme la résistance à la verse ou aux stress biotiques et abiotiques.

Il est connu que les lignines des parois secondaires sont le frein principal à l'accessibilité des sucres de structure et leur quantité réduit la digestibilité de la biomasse (Ermawar *et al.*, 2015, Buxton et Redfearn 1997; Wilson et Hatfield, 1997). La quantité de lignines n'est pas la seule en jeu et sa structure, notamment la composition en différents monolignols et les types de liaisons, ses liens avec les acides hydroxycinnamiques et les autres structures des parois, sont des paramètres importants de la digestibilité des parois (Boerjan *et al.*, 2003). En outre, Le degré de cristallinité de la cellulose et l'ensemble des interactions des composants pariétaux inclus dans la matrice de cellulose peuvent impacter les propriétés de la biomasse (Demura et Ye, 2010). Enfin, même si les pectines sont mineures dans la paroi secondaire elles peuvent réduire significativement la lyse des polysaccharides (Lionetti *et al.*, 2010). Les parois secondaires étant toujours incluses dans une paroi primaire, la complexité de celle-ci et sa résistance à la saccharification via par exemple les liens formés par les pectines, peut impacter la saccharification de la paroi secondaire.

Plusieurs groupes de caractères d'intérêts liés à la qualité de la biomasse font l'objet d'études. Ce sont des caractères de composition, de digestibilité et de saccharification des parois cellulaires et du métabolisme des sucres solubles.

11.1. Evaluation de la qualité de la biomasse

Les caractéristiques de composition de la tige sont déterminées par des mesures histochimiques qui apportent des informations spatiales et des mesures chimiques qui peuvent servir de référence pour l'obtention de courbes de prédiction acquises par spectrométrie proche infra-rouge (NIRS) utilisées pour des estimations indirectes.

La composition est déterminée par plusieurs indicateurs comme la teneur en Matière Minérale (MM) qui représente les résidus inorganiques après calcination et la teneur en Matière Azotée

Totale qui représente la teneur en protéines et est obtenue par la méthode de Kjeldhal (Sáez-Plaza *et al.*, 2013).

Différentes mesures des teneurs spécifiques en composants pariétaux existent. Ces quantifications reposent sur des méthodes chimiques directes et sur des traitements progressifs de la matière sèche qui dépendent de tous les éléments pariétaux. Par exemple, les méthodes de Van Soest *et al.* (1991) pour obtenir les Neutral Detergent Fiber (NDF), Acid Detergent Fiber (ADF) et Acid Detergent Lignin (ADL) permettent de déterminer les proportions des différentes fractions dans les fibres de la tige (Figure I-17). La fraction de fibres insolubles dans les détergents neutres (NDF) correspond à la fraction organique des fibres. Elle contient tous les composants des parois hormis certaines pectines. La fraction insoluble dans les détergents acides (ADF) correspond à la fraction la moins digestible et est composée de cellulose et de lignine. Elle est un indicateur indirect de la digestibilité. La fraction ADL est composée de lignines et est, elle, la fraction des fibres totalement non digestibles. La mesure de ces 3 fractions permet d'estimer des teneurs en cellulose et hémicellulose et de les considérer relativement au contenu total en paroi (NDF). Une autre mesure de la teneur de Cellulose Brute de Weende (CBW) existe également (Thera, 2017).

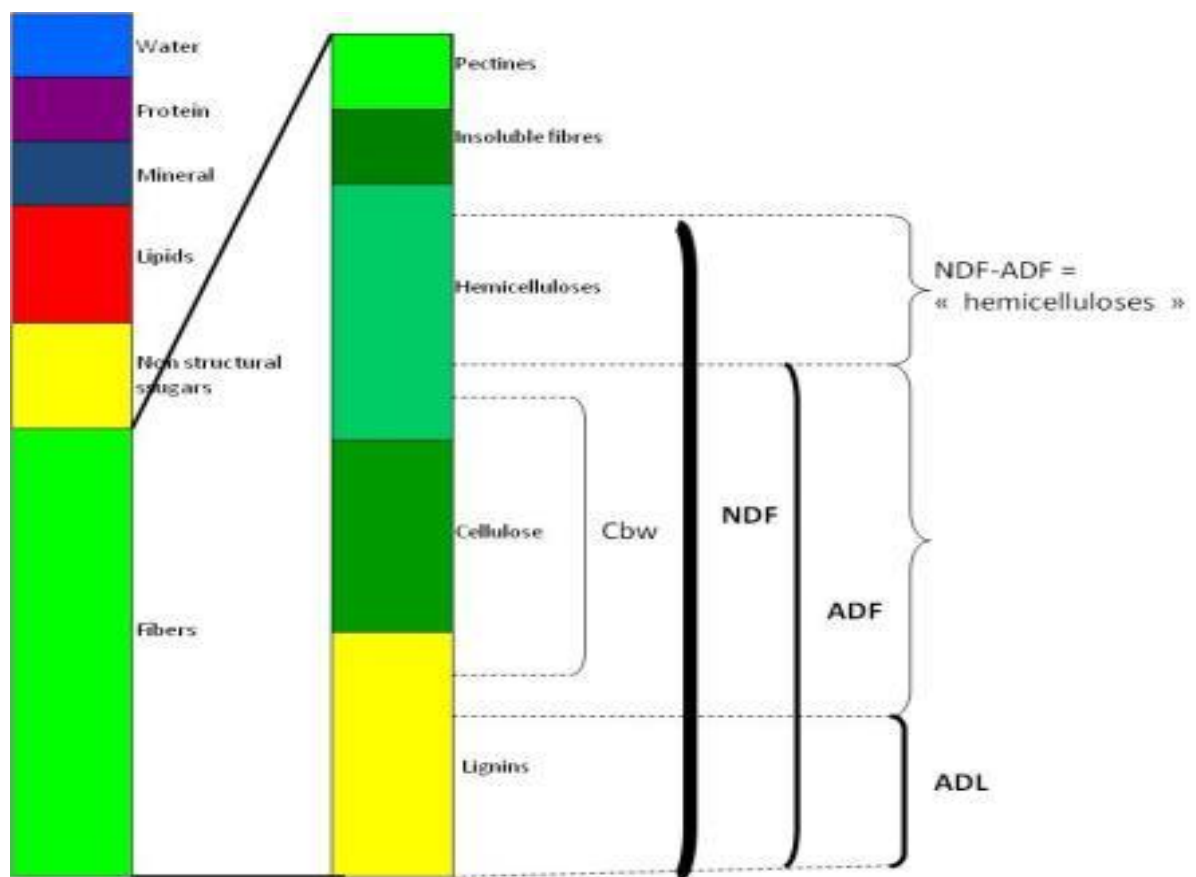


Figure I-17 : Représentation de la portion des composés pariétaux capturée par les mesures de Van Soest et CBW

Des mesures intégratives de la qualité de la biomasse peuvent être réalisées. La solubilité, ou digestibilité, de la biomasse est déterminée par la méthode de Aufrère *et al.* (2007). Cette méthode est basée sur une digestion enzymatique (pepsine-cellulase) qui permet de prédire la Solubilité de la Matière Organique in vitro (SMO) et la Solubilité de la Matière Sèche (SMS). Grâce à ces mesures il est possible de connaître la digestibilité des parois, de la fraction NDF (IVNDFD, In-Vitro NDF Digestibility) qui suppose que la fraction non NDF est totalement digestible.

$$IVNDFD = \frac{(SMO - ((100 - MM) - NDF))}{NDF} \times 100$$

Les outils de mesure de la qualité de la biomasse sont très importants pour les programmes de sélection et les expérimentations qui cherchent à comprendre les déterminants génétiques et moléculaires de la qualité de la biomasse et de la composition des parois secondaires du sorgho.

11.2. Déterminisme génétique de la qualité de la biomasse

Le déterminisme génétique des caractères de qualité de la biomasse ligno-cellulosique a fait l'objet de peu d'études chez le sorgho que ce soit pour la composition des parois secondaires que pour ses propriétés de dégradabilité.

Plusieurs gènes majeurs impactant la qualité de la biomasse ligno-cellulosique ont été identifiés chez le sorgho via l'analyse de mutants présentant un contenu en lignines réduit, une saccharification améliorée et une meilleure digestibilité (Palmer *et al.*, 2008; Xin *et al.*, 2008; Saballos *et al.*, 2008, 2009, 2012, Petti *et al.*, 2013, Sattler *et al.*, 2014, Vermerris, 2011, Vermerris *et al.*, 2007). Il s'agit des mutants *brown midrib (bmr)* et *red for green (rg)* qui présentent des teneurs en lignines modifiées. Les loci mutés des plantes *bmr* incluent une 4CL (*bmr2*), une CAD (*bmr6*) et une COMT (*bmr12* et *bmr8*) qui sont des enzymes de la voie de biosynthèse des monolignols (Saballos *et al.*, 2009, Sattler *et al.*, 2009, Scully *et al.*, 2016b). Plusieurs d'entre eux (*bmr6* et *bmr12*) ont été mobilisés pour le développement de lignées à faible teneur en lignine destinées à l'utilisation sous forme de fourrage (Pedersen *et al.*, 2008). Les mutations responsables d'autres phénotypes Bmr non alléliques aux gènes déjà clonés restent à caractériser (Sattler *et al.*, 2014).

Si ces gènes constituent de formidables outils de contrôle des teneurs en lignines dans la biomasse, leur utilisation en sélection a souvent été contrariée par des verses importantes des

génotypes portant les allèles mutés. Il est aussi important de noter ici, que la variabilité génétique observée pour les teneurs en lignines ne dépend pas uniquement de ces gènes majeurs et que la digestibilité de la biomasse et plus largement ses propriétés structurales ne dépendent pas uniquement des teneurs en lignines mais aussi, comme mentionné précédemment, des proportions de celluloses, hémicelluloses et pectines dans les parois et de leurs interactions avec les lignines (Murray, Rooney, *et al.*, 2008; Shiringani et Friedt, 2011; Trouche *et al.*, 2014). Une meilleure identification des régions chromosomiques responsables du contrôle de la variabilité des différents éléments composant la biomasse et ses propriétés est donc requise pour optimiser la sélection de variétés adaptées à des utilisations spécifiques.

Plusieurs études (Murray *et al.*, 2008, Shiringani et Friedt, 2011) portant sur des populations biparentales basées sur des croisements entre des sorghos grains et des sorghos sucriers ont montré que les régions chromosomiques associées à la teneur en carbohydrates structuraux et en lignines colocalisent largement avec celles associées à la croissance en hauteur et à la floraison. Même si plusieurs régions communes entre ces études ont été détectées, un effet environnemental fort conduit à une instabilité des régions génétiques identifiées. Murray *et al.* (2008) ont également montré que des déterminismes indépendants existent pour ces caractères entre les feuilles et la tige. Ce dernier constat est partagé par Vandenbrink *et al.* (2013) qui s'est intéressé au potentiel d'hydrolyse de la biomasse lignocellulosique et à l'indice de cristallinité de la cellulose (qui confère de la stabilité face à l'hydrolyse) au contraire de Murray *et al.* (2008) et Shiringani et Friedt (2011) qui étaient focalisés sur les caractères descriptifs de composition de la biomasse (NDF, ADF, ADL).

L'exploration d'une diversité génétique plus large via l'analyse d'un panel d'individus non apparentés a été entreprise par Brenton *et al.* (2016) qui se sont intéressés à la teneur en biomasse des parois au travers du caractère de la teneur en fibre. Si 5 régions chromosomiques indépendantes ont été identifiées, il n'y a pas eu de comparaisons avec les régions détectées dans les références mentionnées ci-dessus.

Plusieurs études ont eu pour ambition d'identifier les gènes acteurs de la biosynthèse des carbohydrates structuraux et des lignines dans les tiges (Shakoor *et al.*, 2014; McKinley *et al.*, 2016; Rai *et al.*, 2016). Ces études ont permis de compiler des listes de gènes de structure potentiellement impliqués dans la biosynthèse des parois chez le sorgho. Au-delà des gènes impliqués dans la biosynthèse et la régulation des lignines, d'autres études se sont focalisées aux familles de gènes de structure dans le contexte de la biosynthèse des polysaccharides. Les CESA et CSL (Cellulose-Synthase-like) ont été étudiées chez le sorgho (Rai *et al.*, 2016). En

particulier, 12 CESA et 36/37 CSL ont été identifiées (Paterson *et al.*, 2009; Yin *et al.*, 2009). Les familles CSLF et CSLH ont également fait l'objet de caractérisations (Ermawar *et al.*, 2015). La modification du contenu en hémicelluloses (xylane et MLG) a aussi été explorée (Biswal *et al.*, 2015, Vega-Sanchez *et al.*, 2015).

Néanmoins, peu de liens solides ont été mis en évidence entre les gènes différentiellement exprimés lors de l'évolution de la composition de la biomasse et les régions chromosomiques contribuant à la variabilité génétique des différentes composantes de la biomasse. En conclusion, malgré l'importance du pilotage de la composition de la biomasse pour les applications classiques (alimentation animale basée sur le fourrage ou l'ensilage) et les applications émergentes (production d'énergies renouvelables et de biomatériaux) une faible compréhension du déterminisme génétique des composantes de la biomasse est actuellement disponible. En complément des gènes à effet majeurs (gènes Bmr) qui induisent souvent des effets indésirables, quelques régions à effets relativement faibles ont été mises en évidence, mais une très large part de la variabilité génétique disponible est encore inexploitée.

Dans ce contexte, l'identification des acteurs génétiques impliqués dans le contrôle de la composition des parois secondaires et plus largement de la qualité de la biomasse reste un enjeu important. Si les gènes de structure ont fait l'objet d'études, sans conclure fermement sur leur rôle dans le contrôle de la variabilité phénotypique chez le sorgho, très peu de recherches se sont intéressées aux Facteurs de Transcription (FT) responsables de la régulation de la mise en place des parois secondaires chez cette espèce.

12. Facteurs de transcription impliqués dans la régulation de la mise en place des parois secondaires

De nombreuses familles de FT comme bHLH (Bosch *et al.*, 2011), HB (Cassan-Wang *et al.*, 2013), bZIP (Leene *et al.*, 2016), KNOX (Townsend *et al.*, 2013, Wuddineh *et al.*, 2016, Li *et al.*, 2012), BEL et BEL-like (BLH) (Liu *et al.*, 2014), OFP (Liu et Douglas, 2015, Schmitz *et al.*, 2015), C2H2 (Hirano *et al.*, 2013a, Huang *et al.*, 2017), WRKY (Wang *et al.*, 2010, Yang *et al.*, 2016) et AP2/ERF (Aharoni *et al.*, 2004) sont ou pourraient être impliquées dans la régulation de la mise en place des parois. Cependant, les familles MYB (orthologues chez les plantes des oncogènes Myeloblastosis) et NAC (NAM, ATAF, CUC) sont particulièrement impliquées dans la régulation de la biosynthèse des parois secondaires. Les connaissances au sujet de ces deux familles sont les plus avancées chez *Arabidopsis thaliana* ainsi que chez d'autres dicotylédones comme le peuplier ou l'eucalyptus et se développent de plus en plus

chez les monocotylédones avec quelques études qui ciblent le maïs, le riz et brachypodium (Dixon et Rao, 2018, Wang et Dixon 2012). Comme mentionné auparavant, encore peu d'études s'intéressent aux FT NAC et MYB dans le contexte des parois chez le sorgho. Cependant, deux NAC situés sur des locus associés aux caractères de composition de la biomasse ont été identifiés. Un gène majeur contrôlant la composition de la biomasse causant des tiges sèches et non juteuses a été mis en évidence à travers une analyse de génétique d'association impliquant un panel de diversité ciblant les idéotypes de sorgho sucré et biomasse par Burks *et al.* (2015). Le gène NAC74a, orthologue du gène d'*Arabidopsis* AtNAC074 qui n'est pas caractérisé fonctionnellement chez cette espèce, a été mis à jour par une cartographie fine et semble être responsable de l'association observée (Xia *et al.*, 2018, Zhang *et al.*, 2018). Un orthologue sorgho de NST1 se trouve également au niveau d'une zone chromosomique associée à la variabilité de la saccharification, sans que des preuves supplémentaires de son impact n'aient été apportées (Wang *et al.*, 2013). De récentes études ont également validé fonctionnellement chez le sorgho le rôle de SbMYB60 qui joue un rôle dans la régulation du dépôt des lignines (Scully *et al.*, 2016a, 2017).

12.1. Facteurs de Transcription NAC

12.1.1. Caractéristiques de la famille NAC

Les FT de la famille des NAC forment une des plus grandes familles de FT chez les plantes. Les protéines NAC possèdent généralement un domaine conservé NAM en partie N-terminal et une partie plus divergente destinée à la régulation spécifique en C-terminal. Le domaine NAC consiste en environ 160 acides aminés divisés en 5 sous-domaines de A à E (Figure I-18). Les sous-domaines A, C et D sont communément très conservés alors que les sous-domaines B et E sont très divergeants et confèrent certainement une diversité fonctionnelle aux FT NAC. Le sous-domaine A jouerait un rôle important dans la formation de dimères fonctionnels avec d'autres protéines NAC tandis que les sous-domaines C et D sont capables de se lier à l'ADN (Ooka *et al.*, 2003).

Les gènes NAC ont des rôles divers dans la régulation de plusieurs processus développementaux comme la formation des organes, des embryons et des fleurs (Aida *et al.*, 1997, Souer *et al.*, 1996, Sablowski *et al.*, 1998), le développement de racines (Hibara *et al.* 2003), du méristème apical (Aida *et al.*, 1997, Souer *et al.*, 1996), la sénescence des feuilles (Lee *et al.*, 2012), cycle cellulaire (Kim *et al.*, 2006, Willemsen *et al.*, 2008), la signalisation hormonale (Kim *et al.*, 2006, Fujita *et al.*, 2004) et la réponse aux stress abiotiques (Puranik,

et al., 2012, He *et al.*, 2005, Tran *et al.*, 2004, Balazadeh *et al.*, 2010, Guan *et al.*, 2014) et biotiques (Nuruzzaman *et al.*, 2013).

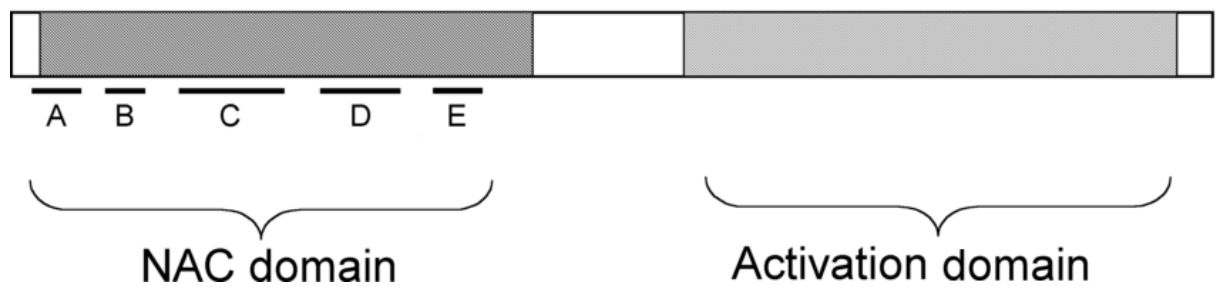


Figure I-18 : Structure typique des protéines de la famille des facteurs de transcription NAC (D'après Ooka *et al.* 2004)

12.1.2. Rôle des NAC dans la régulation des parois secondaires

Le premier niveau de régulation de la biosynthèse des parois secondaires est constitué chez *Arabidopsis* de 11 gènes de la famille des NAC responsables du déclenchement du dépôt des parois secondaires (Wang et Dixon, 2012, Zhong *et al.*, 2006; Mitsuda *et al.*, 2007) (Figure I-20). Chez cette espèce, différents types de NAC sont responsables de ce dépôt dans les cellules vasculaires et dans les cellules des fibres.

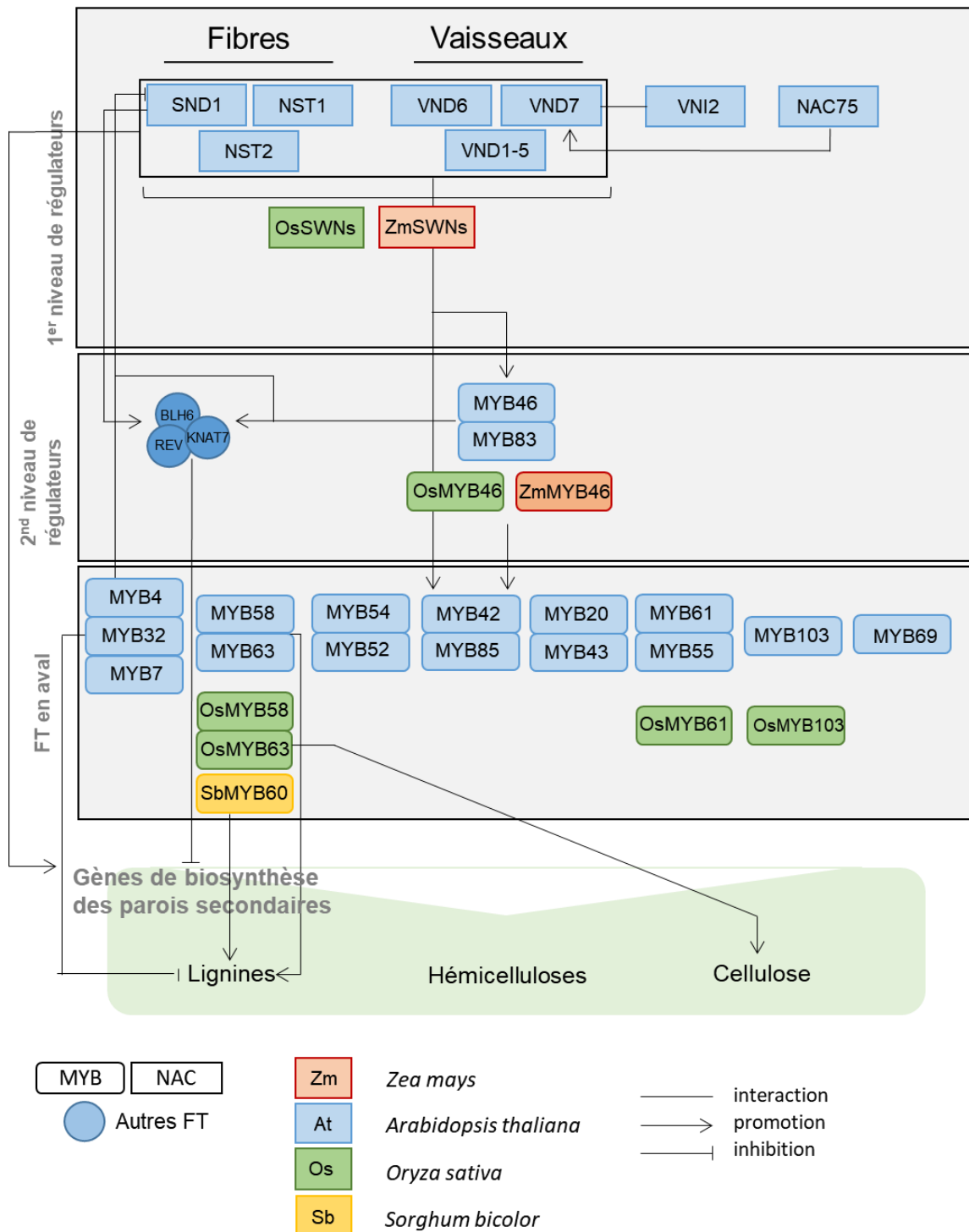


Figure I-20 : Facteurs de Transcription MYB et NAC et autres régulateurs des parois secondaires validés chez Arabidopsis, le maïs, le riz et le sorgho.

Les gènes VASCULAR-RELATED NAC-DOMAIN (VND), qui sont au nombre de 7 chez *Arabidopsis* (Shen *et al.*, 2009), sont responsables de la différenciation des éléments vasculaires (proto et méta xylèmes). Ils contrôlent le développement des parois secondaires et la mort programmée des cellules vasculaires (Kubo *et al.*, 2005; Yamaguchi *et al.*, 2008).

Les gènes NAC SECONDARY WALL TICKENING PROMOTING FACTOR et SECONDARY WALL-ASSOCIATED NAC DOMAIN (NST et SND) promeuvent le dépôt des parois secondaires dans les fibres de support entre les vaisseaux conducteurs de sève. Ils interviennent, au contraire des gènes VND, après l'acquisition de l'identité des cellules. Chez *Arabidopsis*, NST1 et NST3/SND1 agissent de façon redondante dans les cellules de la tige tandis que NST2 est spécifiquement exprimé dans les anthères et régule de façon redondante avec NST1 la régulation des parois secondaires dans ce tissu. Tous ces gènes sont très proches phylogénétiquement (Shen *et al.*, 2009) et VND6 et VND7 sont même capables de compléter le phénotype de mutants *nst1snd1* (Zhong *et al.*, 2010a). SND2 et SND3, de proches homologues des gènes cités plus haut, sont des acteurs avals de la régulation et sont activés par les autres gènes des familles VND et NST (Zhong *et al.*, 2008).

La conservation des fonctions de ces gènes a été observée chez des graminées entre le riz, le maïs, *brachypodium* et *switchgrass*. La surexpression hétérologue des orthologues riz et maïs complète le phénotype des mutants *Arabidopsis nst1snd1*. Certaines différences d'expression de ces gènes, appelés SWN (Secondary Cell Wall NACs) chez les graminées, ont été identifiées. En effet, toutes les classes de NAC mentionnées ci-dessus montrent un même patron d'expression chez les graminées alors que les gènes VND sont spécifiques des vaisseaux et les NST sont spécifiques des fibres chez *Arabidopsis* (Zhong *et al.*, 2011, 2015; Valdivia *et al.*, 2013; Yoshida *et al.*, 2013; Xiao *et al.*, 2018).

Deux autres gènes de la famille des NAC sont également impliqués dans la régulation des parois secondaires, il s'agit de VNI2 (VND INTERACTING 2) (Yamaguchi *et al.*, 2010) et de NAC075 (Endo *et al.*, 2015).

12.2. Facteurs de transcription MYB

12.2.1. Caractéristiques de la famille MYB

La famille des MYB est également une très grande famille de FT fonctionnellement diverse. Cette famille est caractérisée par un domaine spécifique (MYB-domain) qui comporte une cinquantaine d'acides aminés répétés de façon imparfaite jusqu'à 4 fois, chaque répétition

s'enroulant en trois hélices alpha (Figure I-19). Cet ensemble d'hélices forme un cœur hydrophobe et la 3^{ème} hélice de chaque répétition est responsable de la reconnaissance de l'ADN en s'intercalant dans le grand sillon. Le type de motifs d'acides aminés répétés détermine la classe à laquelle appartient la protéine MYB par rapport aux 3 répétitions (R1, R2, R3) présentes dans la protéine originelle c-MYB dont est issue la famille (Dubos *et al.*, 2010).

La plupart des gènes MYB chez les plantes (une centaine selon les espèces) appartient à la classe des R2R3-MYB. La séquence des gènes de la classe des R2R3-MYB est très variée au-delà du domaine très conservé MYB avec un domaine d'activation et de répression. Plusieurs études de génomique comparative ont mené à l'identification de sous-groupes de R2R3-MYB qui ne sont pas présents chez *Arabidopsis*, suggérant une grande spécialisation de ces groupes chez les espèces qui les ont conservés ou acquis après la spéciation. L'expansion des R2R3-MYB chez les plantes (Dubos *et al.*, 2010) est cohérent avec les rôles spécifiques mis en évidence dans le métabolisme des plantes.

Au-delà des R2R3-MYB qui contiennent la majorité des gènes impliqués dans la mise en place des parois cellulaires, d'autres classes de MYB existent. La plus petite classe (4R-MYB pour 4 répétitions R1/R2) est peu présente chez les plantes. Les gènes 3R-MYB (R1R2R3-MYB) dont seraient issus les R2R3-MYB, sont typiquement au nombre de cinq chez les plantes supérieures et responsables du contrôle du cycle cellulaire (Rosinski, 1998, Ito, 2005, Haga, 2007). Les « MYB-related » sont le 4^{ème} groupe de MYB et sont des protéines diverses présentant un seul domaine MYB qui peut être partiel. Elles régulent la morphogénèse cellulaire (Pesch et Hulskamp, 2009, Simon *et al.*, 2007), le métabolisme secondaire (Dubos *et al.*, 2008, Matsui *et al.*, 2008), l'horloge circadienne (Lu *et al.*, 2009), l'organogénèse (Kerstetter *et al.*, 2001), le développement des chloroplastes (Waters *et al.*, 2009) et la réponse au manque de phosphate (Rubio *et al.*, 2001).

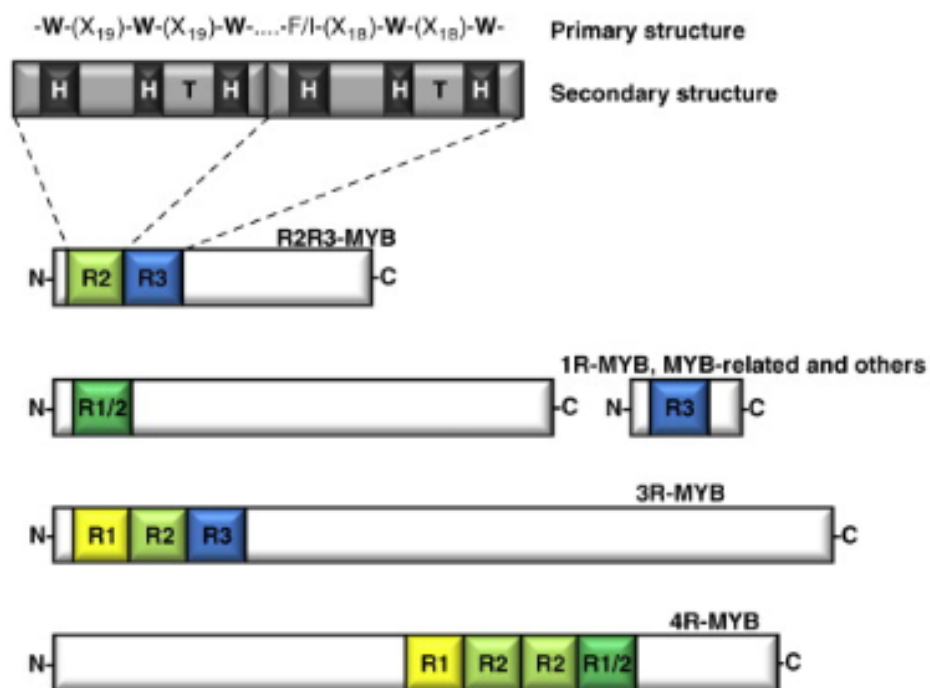


Figure I-19 : Structures typiques des protéines de la famille des facteurs de transcription MYB et les structures spécifiques aux 4 classes de MYB (D'après Dubos *et al.*, 2010)

12.2.2. Rôle des MYB dans la régulation des parois secondaires

Le deuxième niveau de régulation des parois secondaires ainsi que la régulation en aval sont assurés en majorité par des gènes de la famille des MYB (Zhong *et al.*, 2007, Ko *et al.*, 2009, McCarthy *et al.*, 2009) (Figure I-20). En particulier, chez *Arabidopsis*, MYB46 et son paralogue MYB83 coordonnent de façon redondante l'activation d'une pléthore de FT MYB « avals » ainsi que de gènes de structures responsables de la biosynthèse des composés pariétaux. Leurs surexpressions activent les gènes responsables de la biosynthèse de la cellulose, des hémicelluloses et des lignines ainsi qu'un dépôt ectopique de parois secondaires. AtMYB46 est également une cible directe des FT NAC du premier niveau de régulation (VND6, VND7, SND1). Ceci est cohérent avec l'observation du phénotype drastique des doubles mutants *myb46/myb83* qui présentent un phénotype de réduction des parois secondaires plus sévère par rapport aux mutants *nst1/snd1*. En effet, MYB46 et MYB83 ne sont pas spécifiques des seuls tissus fibreux mais agissent aussi dans les tissus vasculaires (Zhong *et al.*, 2007, McCarthy *et al.*, 2009, Ohashi-Ito *et al.*, 2010; Yamaguchi *et al.*, 2011).

De nombreux régulateurs avals sont induits par l'expression de MYB46 et MYB83 chez *Arabidopsis*. Parmi eux se trouvent MYB43, MYB52, MYB54, MYB58, MYB63, MYB4 et MYB32 (Zhong *et al.*, 2008, 2015; Zhou *et al.*, 2009). Certains d'entre eux sont également directement reconnus par SND1 et en retour régulent son expression négativement. SND1 régule également de nombreux MYB : MYB103, MYB85, MYB52, MYB54, MYB69, MYB42, MYB43 et MYB20. MYB103 pourrait être une cible directe de SND1 (Zhong *et al.*, 2008) mais aussi de VND6 et VND7 (Ohashi-Ito *et al.*, 2010, Yamaguchi *et al.*, 2011). MYB58, MYB63, MYB42 et MYB85 sont spécifiquement impliqués dans la régulation des lignines (Zhong *et al.*, 2008; Zhou *et al.*, 2009, Zhao et Dixon, 2011).

Certains MYB sont des régulateurs négatifs des parois, il s'agit de MYB4, MYB7 et MYB32 (Zhao et Bartley, 2014). L'étude de Zhao et Bartley (2014) suggère que MYB4 serait responsable du détournement de la voie des phénylpropanoïdes vers d'autres voies métaboliques. En effet MYB4 réprime l'expression des gènes C4H et 4CL, codant pour des enzymes très en amont de la voie des phénylpropanoïdes (Jin *et al.*, 2000). Cependant MYB7 et MYB32 répriment et induisent, respectivement, des gènes de la voie des flavonoïdes (Preston *et al.*, 2004; Fornalé *et al.*, 2014), suggérant que les cibles de ces deux MYB sont différentes.

Chez les graminées, de plus en plus de FT MYB impliqués dans la mise en place des parois secondaires sont identifiés. ZmMYB46 et OsMYB46 peuvent compléter les défauts de parois secondaires des mutants *Arabidopsis myb46/myb83* (Zhong *et al.*, 2011, 2015). De plus leur surexpression mène au dépôt ectopique de parois secondaires dans la tige et à une augmentation des teneurs en cellulose, xylanes et lignines (Zhong *et al.*, 2011, 2015, Ko *et al.*, 2012, Kim *et al.*, 2013). L'orthologue sorgho SbMYB60 de AtMYB58 et AtMYB63 semble partager avec ces derniers le rôle de régulateur des lignines (Scully *et al.*, 2016a, 2017). Les orthologues riz des 4 MYB régulateurs des lignines chez *Arabidopsis* (OsMYB58/63 et OsMYB42/85) montrent les mêmes caractéristiques (Hirano *et al.*, 2013b). OsMYB103 a également été fonctionnellement validé pour son rôle dans la promotion du dépôt des parois secondaires (Hirano *et al.*, 2013b, Yang *et al.*, 2014, Ye *et al.*, 2015). Les homologues chez le maïs des répresseurs de la paroi secondaire, ZmMYB11, ZmMYB31 et ZmMYB52 réguleraient négativement les gènes COMT et 4CL même si certaines différences ont été mises en évidence (Vélez-Bermúdez *et al.*, 2015)

12.3. Autres régulateurs des parois secondaires

Quelques FT n'appartenant pas aux familles des MYB et des NAC jouent des rôles importants dans la régulation des parois secondaires. Ce sont les FT WRKY12, KNAT7 et BLH6 (Figure I-20).

Chez *Arabidopsis*, le peuplier et la luzerne (*Medicago sativa*), la réduction de l'expression de WRKY12 génère une augmentation du dépôt des parois secondaires dans le parenchyme de la tige (Wang *et al.*, 2010; Gallego-Giraldo *et al.*, 2016; Yang *et al.*, 2016). Chez le maïs et switchgrass, des altérations de la mise en place des parois secondaires sont observées sur des lignées où WRKY12 est réprimé. WRKY12 serait donc un répresseur des parois secondaires chez les dicotylédones et les monocotylédones. Des interactions entre ce gène et le promoteur de NST2 ainsi qu'une auto-répression sont observées chez la luzerne et *Arabidopsis* (Wang *et al.*, 2010).

Chez *Arabidopsis*, KNAT7 et BLH6 sont également considérés comme des répresseurs de la biosynthèse des parois secondaires (Li E. *et al.*, 2012; Liu *et al.*, 2014). L'expression de KNAT7 peut être induite par MYB46/83 et MYB61 qui sont des activateurs des parois secondaires (Zhong *et al.*, 2008; Romano *et al.*, 2012). Ceci suggère l'existence d'une boucle de rétrocontrôle négative pour le contrôle de la biosynthèse des parois secondaires (Li E. *et al.*, 2012; Liu et Douglas, 2015).

13.Objectifs de la thèse et organisation du manuscrit

Ce travail de thèse vise à identifier et comprendre l'implication de gènes dans la mise en place des parois secondaires et les FT responsables du dépôt de celles-ci. La première partie (chapitre II) de la thèse porte sur l'identification de gènes candidats tandis que la deuxième partie (chapitre III) porte sur la validation fonctionnelle de ces gènes par des approches de génétique inverse.

L'identification de gènes candidats a été faite en utilisant plusieurs approches. Une analyse phylogénétique des principales familles de FT des familles NAC et MYB a été menée afin d'identifier, à partir des données de la littérature, quels sont les gènes potentiellement impliqués dans la régulation des parois secondaires chez le sorgho. Une meilleure compréhension des liens phylogénétiques de ces familles entre 9 espèces de monocotylédones et dicotylédones permet également d'identifier des clades spécifiques au lignage monocotylédones et d'avoir une vision d'ensemble des acteurs potentiels. L'utilisation d'études d'association entre génotype et phénotype liés à la composition de la biomasse des tiges a également permis de dresser des listes de gènes candidats dont la variation allélique entraîne possiblement des variations phénotypiques liées à la composition de la tige et des parois secondaires. Cette approche a mené à l'identification de gènes candidats à travers toutes les fonctions protéiques et non seulement les FTs. Enfin, une autre approche mêlant une approche de bioinformatique et des données transcriptomiques en construisant des réseaux de co-expression de gènes a permis d'identifier de nouveaux gènes candidats, de renforcer nos connaissances de la spatio-temporalité d'expression de ces gènes et également de mettre à jour de potentielles interactions entre les gènes mis en évidence. Une publication rassemblant les aspects de phylogénie et de réseaux de co-expression de gènes est acceptée avec révisions mineures dans le journal *Frontiers in Plant Science*.

La deuxième partie de la thèse s'intéresse à la validation fonctionnelle des gènes candidats identifiés en première partie, en particulier les FT. Plusieurs approches ont été testées. La première repose sur la transformation du sorgho par des versions modifiées des gènes candidats afin de mettre en évidence le rôle de ces modifications sur le phénotype de plantes mutantes. La deuxième repose également sur une approche « mutant » mais n'a pas été menée en transformation stable mais en système transitoire avec utilisation de protoplastes de sorgho. Ces expérimentations ont fait l'objet de nombreuses étapes d'optimisation et restent à maîtriser complètement.

Chapitre II : Identification de gènes candidats impliqués dans la détermination de la qualité de la biomasse chez le sorgho

1. Introduction

Les recherches sur *Arabidopsis thaliana* ont mené à une bonne connaissance des Facteurs de Transcription (FT) impliqués dans la régulation des parois secondaires. La plupart de ces gènes sont membres des familles de FT MYB et NAC (Zhong *et al.*, 2008). Ils agissent à différents niveaux et occupent des rôles de régulateurs principaux comme de régulateurs plus spécifiques de la synthèse de certains composés pariétaux. La fonction de certains des régulateurs principaux est majoritairement conservée chez différentes espèces dont le riz, maïs, peuplier et eucalyptus (Goicoechea *et al.*, 2005; McCarthy *et al.*, 2009; Zhong *et al.*, 2011, 2013). Ce socle de connaissances des FT développé chez *Arabidopsis* et quelques autres espèces modèles constitue une opportunité d'accélérer et de faciliter la découverte des gènes impliqués dans la régulation des parois secondaires chez le sorgho en particulier et les graminées en général. Chez le sorgho, la compréhension des mécanismes de la mise en place des parois secondaires est très partielle avec un seul FT (SbMYB60) capable d'induire la synthèse de lignines validé fonctionnellement jusqu'à présent (Scully *et al.*, 2016b). En plus de ces travaux de génétique fonctionnelle, deux études de génétique d'association ont mis en évidence l'impact potentiel d'un homologue sorgho de NST1 (Wang *et al.*, 2013a) et de SbNAC074 (Casto *et al.*, 2018; Xia *et al.*, 2018; Zhang *et al.*, 2018), dont un homologue *Arabidopsis* est annoté comme étant impliqué dans la mort cellulaire, sur la qualité de la biomasse et la composition des parois secondaires.

En dehors des FT, seuls quelques gènes de structure majeurs impactant la qualité de la biomasse ligno-cellulosique ont été validés chez le sorgho via l'analyse de mutants présentant un contenu en lignines réduit, une saccharification améliorée et une meilleure digestibilité (Palmer *et al.*, 2008; Xin *et al.*, 2008; Saballos *et al.*, 2008, 2009, 2012, Petti *et al.*, 2013, Sattler *et al.*, 2014, Vermerris, 2011, Vermerris *et al.*, 2007). Il s'agit notamment des mutants *bmr* qui présentent des teneurs en lignines modifiées. Les loci mutés des plantes *bmr* incluent une 4CL (*bmr2*), une CAD (*bmr6*) et une COMT (*bmr12* et *bmr8*) qui sont des enzymes de la voie des monolignols (Saballos *et al.*, 2009, Sattler *et al.*, 2009, Scully *et al.*, 2016b)

L'objectif des analyses décrites dans ce chapitre est donc de contribuer à une meilleure compréhension du développement et de la régulation de la composition de la biomasse des entre-nœuds chez le sorgho avec un focus particulier sur les FT MYB et NAC. Afin d'atteindre cet objectif nous avons combiné plusieurs approches permettant d'identifier des gènes candidats impliqués dans la mise en place des parois secondaires. Dans un premier

temps, les homologues entre gènes déjà validés pour leur rôle dans la régulation des parois secondaires chez différentes espèces et les gènes présents chez le sorgho ont été explorées grâce à la mise à jour des phylogénies des familles de FT MYB et NAC chez 9 espèces différentes. Dans un deuxième temps, nous avons réalisé des analyses de réseau de co-expression fondées sur l'échantillonnage d'entre-nœuds de sorgho en développement aux cours de 3 années d'expérimentations au champ. Cette méthode, dont la force d'analyse est démontrée depuis quelques années, permet d'explorer des données haut débit et de disséquer des processus biologiques encore mal connus (Higashi et Saito, 2013) comme la mise en place des parois secondaires chez *Arabidopsis* et différentes espèces de céréales (Hirano *et al.*, 2013b, 2013a; Hu *et al.*, 2017; Ruprecht *et al.*, 2011; Sibout *et al.*, 2017). Ces analyses permettront de mettre en évidence, non seulement le rôle des FT mais également des gènes de structure dont les contributions à ces réseaux seront mises en évidence.

Les résultats des deux premières parties de ce chapitre ont fait l'objet d'une publication (Hennet *et al.*, 2020). Les résultats qui ne pourront être représentés par des figures et tableaux dans ce manuscrit, à cause de contraintes de taille et de nature des fichiers, sont décrits en matériels supplémentaires de cette publication.

1.1. Analyses phylogénétiques

Les connaissances fonctionnelles acquises chez certaines espèces peuvent être mobilisées pour l'inférence de fonctions chez d'autres espèces par la mise en place d'analyses phylogénétiques. Dans le cadre de la mise en place des parois secondaires, un socle de connaissances solide existe chez *Arabidopsis* concernant les FT MYB et NAC. Ces familles ont d'ailleurs fait l'objets de revues détaillées (Dubos *et al.*, 2010; Ooka *et al.*, 2003). Les connaissances se développent de plus en plus chez d'autres dicotylédones et monocotylédones mais restent rares chez le sorgho. Une analyse phylogénétique globale de ces familles chez plusieurs espèces permettra donc de faire le lien entre les connaissances acquises chez ces différentes espèces et d'émettre des hypothèses concernant le sorgho. Une vision d'ensemble des régulateurs potentiels MYB et NAC de la mise en place des parois secondaires sera également acquise.

1.1.1. Intérêt des analyses phylogénétiques

La connaissance des relations au sein de familles de gènes joue un rôle important dans la compréhension des réseaux de gènes et des systèmes de régulation complexes au cœur

desquels elles sont intégrées. Les analyses phylogénétiques permettent de dévoiler une part de ces relations en apportant des informations sur l'histoire évolutive des gènes et les homologies qui les lient. Les analyses phylogénétiques sont généralement menées sur un groupe de séquences proches alignées issues du séquençage complet des génomes. Elles reposent sur la similarité des séquences comparées et sur le pourcentage de longueur sur laquelle cette similarité est observée.

Conduire une analyse phylogénique à travers différentes espèces, en incluant les gènes issus de génomes qui ont divergé permet d'identifier les liens qui existent entre ces gènes. Leur conservation, perte ou diversification sont autant d'informations qui permettent de comprendre les spécificités d'action de certains gènes chez un taxon ou de faire l'hypothèse de l'existence de fonctionnalités communes entre gènes de taxons différents.

Les homologies de gènes entre différentes espèces ont deux origines. D'un côté, les gènes orthologues descendent d'un même gène présent chez le dernier ancêtre commun et sont issus d'un événement de spéciation. D'un autre côté, les gènes paralogues sont issus d'un événement de duplication indépendant de la spéciation.

La construction d'arbres issus des alignements de séquences homologues permet de visualiser de façon logique l'arborescence des gènes des familles étudiées. Ils sont construits selon différentes méthodes ; maximum de parcimonie (Maximum Parsimony, MP), distances évolutives entre les séquences (Evolutionary Distance, ED) et maximum de vraisemblance (Maximum Likelihood, ML). La première cherche à construire un arbre des séquences alignées avec le moins de mutations possibles, la deuxième calcule des paires de distance génétique entre les séquences et la dernière calcule l'arbre le plus probable expliquant l'alignement existant des séquences (Yang et Rannala, 2012).

Cette dernière méthode est plus coûteuse que les autres en termes de temps de calcul mais produit les résultats les plus corrects (Schmidt, 2002). La méthode ML est largement utilisée dans de nombreux champs de recherche et cela en routine car elle est statistiquement solide et robuste (Zhou *et al.*, 2018). Ainsi, la méthode ML est celle que nous avons utilisée pour construire deux arbres phylogénétiques des familles NAC et MYB à travers plusieurs espèces de monocotylédones et dicotylédones. L'utilisation d'un grand nombre d'espèce permet d'apporter de la robustesse aux résultats phylogénétiques obtenus. De plus, l'utilisation d'espèces dicotylédones et monocotylédones permet d'analyser la conservation des familles de gènes entre ces deux groupes.

1.1.2. Rôles des NAC et MYB dans la régulation des parois secondaires

Les FT NAC et MYB sont deux grandes familles présentes chez les plantes. Elles sont impliquées dans de nombreux processus biologiques comme la réponse aux stress (Nuruzzaman *et al.*, 2013; Rubio *et al.*, 2001) ou le développement des organes (Hibara *et al.*, 2006; Saha *et al.*, 2016). Elles jouent des rôles clés de régulation de la mise en place des parois secondaires (Rao et Dixon, 2018; Wang et Dixon, 2012). Leurs rôles majeurs dans la mise en place des parois secondaires ont été mis en évidence chez *Arabidopsis*, le maïs, le riz, l'eucalyptus et le peuplier (Figure I-20). Le premier niveau de régulation est constitué des gènes NAC appartenant aux sous-groupes NST et VND qui sont, respectivement, responsables de la mise en place des parois secondaires dans les fibres de la tige et dans les faisceaux vasculaires chez *Arabidopsis* (Mitsuda *et al.*, 2005; Yamaguchi *et al.*, 2011; Zhong *et al.*, 2007, 2010). Leur fonction est globalement conservée chez le riz, le maïs, le peuplier, l'eucalyptus et *Arabidopsis*. Cependant, la spécificité spatiale de l'action des homologues de ces gènes reste débattue (Nakano *et al.*, 2015; Ohtani *et al.*, 2011; Zhong, Lee, *et al.*, 2011; Zhong *et al.*, 2010a, 2010b; Zhong, McCarthy, *et al.*, 2011). Le deuxième niveau de régulation est assuré par deux gènes redondants de la famille des MYB, MYB46 et MYB83 dont le rôle est également conservé chez différentes espèces (Kim *et al.*, 2013; Ko *et al.*, 2014; Zhong *et al.*, 2008). Les FT en aval de ces deux niveaux principaux sont également majoritairement des gènes MYB.

Contrairement à *Arabidopsis*, la connaissance des acteurs MYB et NAC impliqués dans la mise en place des parois secondaires est très peu étendue chez le sorgho. Ainsi, des travaux de caractérisation de ces familles sont nécessaires pour permettre des avancées dans la compréhension de la mise en place de la biomasse et des parois chez le sorgho.

1.1.3. Objectifs des analyses phylogénétiques des MYB et des NAC

L'objectif des analyses phylogénétiques présentées dans cette partie est d'obtenir une vision globale des familles MYB et NAC chez le sorgho, chez les monocotylédones en général ainsi que chez des dicotylédones. Les liens phylogénétiques mis en évidence, en particulier avec *Arabidopsis* chez laquelle de nombreuses études de validation fonctionnelles des gènes NAC et MYB ont été menées, constitueront un socle d'analyse pour participer à la détermination de la fonctionnalité des gènes des autres espèces.

1.2. Analyse de réseaux de co-expression de gènes

L'identification de l'ensemble des gènes des deux familles de FT les plus impliquées dans la régulation des parois secondaires chez les plantes permettra de disposer d'une base de données sur laquelle mener des analyses génomiques globales. Ainsi, la construction d'un réseau de co-expression de gènes sur des données transcriptomiques étoffées alliée à la connaissance des clés de voute potentielles de la mise en place des parois permet de participer à leur validation et d'identifier les autres acteurs de ce processus en interrogeant les réseaux de gènes dont ils font partie.

1.2.1. Utilisation des réseaux de co-expression de gènes

L'expression des gènes est un facteur déterminant de la construction du phénotype d'un organisme (Gamazon *et al.*, 2015; Lin *et al.*, 2017). Si l'expression est régulée à différents niveaux, organisation des chromosomes, accessibilité de la chromatine, transcription, transport et modification des ARN, ou modification des protéines, l'étude du transcriptome et des informations portées par les ARN messagers focalise beaucoup de recherches de par sa relative facilité de mise en œuvre. A chacun de ces niveaux, le réseau de régulation s'articule entre régulateurs et cibles. Au niveau transcriptomique, l'identification des régulateurs comme les FTs (FT) et des gènes dont les promoteurs sont les cibles des FT est la clé de la compréhension du réseau régulateur de l'expression génétique (Thompson *et al.*, 2015). Les changements de régulation et d'expression sont des déterminants majeurs de la diversité phénotypique, c'est pourquoi les études comparatives au sein de populations de la même espèce (Liu *et al.*, 2019b; Yu *et al.*, 2017) et d'espèces différentes (Ficklin et Feltus, 2011; Oldham *et al.*, 2006) tirent parti de leurs divergences et similarités qui permettent de mettre au jour les mécanismes de changement d'expression et les réseaux de régulation (Thompson *et al.*, 2015) Wittkopp, 2007.

L'objectif de l'analyse de co-expression de gènes est donc de reconstruire des réseaux de gènes sur la base de leurs similarités d'expression. De telles approches peuvent être menées à partir de différentes sources, en majorité des données de transcriptomiques, telles des micro-arrays (Rao *et al.*, 2017) et le séquençage d'ARN (RNAseq)(Huang *et al.*, 2017). Elles peuvent donc porter sur le transcriptome entier d'un organe ou de différents tissus, voire cellules d'un organisme ((Ma *et al.*, 2018; Takahagi *et al.*, 2018; Tan *et al.*, 2017), un sous-set de gènes déterminés par leur fonctionnalité (Ding *et al.*, 2015; Smita *et al.*, 2015) ou par leur expression différentielle dans des conditions contrastées (Tai *et al.*, 2018).

Des analyses de réseaux de co-expression de gènes aux objectifs variés ont été menées sur différentes monocotylédones proches du sorgho, comme le riz (Hu *et al.*, 2017; Tan *et al.*, 2017), le maïs (Huang *et al.*, 2017; Wang *et al.*, 2018) et brachypodium (Sibout *et al.*, 2017; Sircar *et al.*, 2018). Le sorgho a également fait l'objet de quelques études de ce type dans le cadre d'une étude méthodologique conjointe avec Arabidopsis évaluant un pipeline d'analyse de transcriptome (Proost *et al.*, 2017), d'une base de données transcriptomiques (Tian *et al.*, 2016) et de l'analyse d'un mutant de la voie de biosynthèse de la lignine (Tetreault *et al.*, 2018) ainsi que de la régulation de la formation des parois en ciblant spécifiquement le FT

SbMYB60 (Scully *et al.*, 2017). La mise en place des parois secondaires a également fait l'objet d'études de co-expression de gènes dans le contexte du développement des entre-nœuds chez *Miscanthus* (Hu *et al.*, 2017), la lignification chez *Brachypodium* (Sibout *et al.*, 2017), de la qualité du bois chez l'eucalyptus (Dharanishanthi et Ghosh Dasgupta, 2018), de la compréhension de la mise en place des parois chez le peuplier (Cai *et al.*, 2014) et de l'évaluation d'une méthode d'analyse de réseaux en étudiant le cas de la biosynthèse et la régulation des parois chez le maïs (Huang *et al.*, 2017).

1.2.2. Principe de l'analyse de réseaux de co-expression de gènes

La méthode d'analyse pondérée de co-expression de gènes (Weighted Gene Coexpression Network Analysis, WGCNA) est une des méthodes les plus utilisées pour construire des réseaux de co-expression (Langfelder et Horvath, 2008). C'est la méthode que nous avons choisie d'utiliser ici pour analyser des données RNAseq collectées sur des entre-nœuds de tige de sorgho en développement. Les méthodes de construction de réseau reposant sur la corrélation entre les gènes, comme c'est le cas de WGCNA avec l'utilisation d'un coefficient de corrélation de Pearson, sont préconisées pour l'analyse d'un nombre de gènes supérieur à 10 000 (Krouk *et al.*, 2013).

Toutes les analyses de co-expression de gènes du type de WGCNA sont construites en connectant des paires de gènes dont les profils d'expression sont statistiquement fortement corrélés sur de nombreuses conditions d'échantillonnages (Gupta et Pereira, 2019). En effet, des paires de gènes co-exprimés sont souvent couplés fonctionnellement, par exemple à travers la régulation coordonnée de voies de signalisation/biosynthèse. Identifier et connecter les paires de gènes au niveau du transcriptome entier est ce qui permet de construire le réseau de co-expression qui n'est qu'une façon mathématique et graphique abstraite de représenter les associations qui sont analysées. L'extraction de groupes de gènes dont l'expression est très corrélée permet d'identifier des gènes qui font très probablement partie des mêmes voies de régulation responsables de la synchronisation de leur expression et ont ainsi de grandes chances d'être fonctionnellement reliés ou à défaut d'être impliqués dans des phénomènes synchrones.

1.2.3. Gènes impliqués dans le réseau de régulation des parois chez le sorgho

La mise en place des parois fait intervenir de nombreuses enzymes de biosynthèse, de transport, de modification et de dégradation des éléments pariétaux. Plusieurs études ont utilisées ou générées des listes de gènes potentiellement impliqués dans la biosynthèse des parois. Rai *et al.* (2016) ont utilisé une approche phylogénétique similaire à celle développée dans ce chapitre en partant des familles de gènes impliqués dans la synthèse des parois chez d'autres espèces pour en isoler une signature et pouvoir récupérer les gènes homologues dans le génome du sorgho. D'autres approches ont été utilisées dans cette étude pour interroger les bases de données existantes et identifier de façon la plus exhaustive les gènes potentiellement impliqués dans la mise en place des parois (Rai *et al.*, 2016). Les familles de gènes ciblées dans cette étude sont des gènes impliqués dans la biosynthèse des polysaccharides de structure, leur remodelage et dégradation ainsi que la famille des laccases, impliquée dans la biosynthèse des lignines. Cinq cent vingt gènes de 20 familles ont été identifiés dont des CESA, CSL, PME, XTH, expansines, etc. La majorité de ces enzymes appartiennent à la superfamille des glycosyl transférase qui sont impliquées dans la synthèse de la cellulose, hémicellulose, et des pectines. McKinley *et al.* (McKinley *et al.*, 2016) se sont intéressés à la caractérisation des profils d'expression (différentiels d'expression) des gènes liés aux parois végétales récupérés à partir des annotations d'une base de données et de différentes études (Mitchell *et al.*, 2007; Penning *et al.*, 2009). Ils ont identifié 200 gènes impliqués dans la biosynthèse des parois différentiellement exprimés pendant le développement des entrenœuds de sorgho. Ils ont montré que les profils d'expression des familles de gènes identifiés (GAX, biosynthèse de la cellulose, des lignines, etc.) étaient très variés pendant le développement et, qu'en particulier, les CESA spécifiques des parois secondaires (Tanaka *et al.*, 2003) atteignent un pic d'expression très tardivement dans le développement (7 jours avant l'anthèse). Petti *et al.* ont menés différentes études portant sur la biosynthèse de la cellulose et des lignines chez le sorgho (Petti *et al.*, 2013, 2015). Ils ont permis d'identifier des gènes liés à la biosynthèse de la cellulose dans le cadre de la signalisation des gibbérellines (Petti *et al.*, 2015) et ont également détecté la régulation négative de l'expression des gènes de la biosynthèse des lignines en conséquence d'une mutation entraînant une meilleure saccharification de la biomasse (Petti *et al.*, 2013).

Peu d'études se sont donc intéressées aux gènes liés aux parois sans *a priori* et les connaissances sont encore très incomplètes chez le sorgho. Les liens entre les gènes de

structure responsable de la biosynthèse des parois restent ainsi largement à explorer chez le sorgho. En effet, mis à part Tetreault *et al.* (2018) et Scully *et al.* (2017) qui se sont focalisés sur des transformants bien précis, les réseaux de co-régulations liés à la mise en place des parois n'ont pas été explorés. En particulier, les régulateurs des parois et leurs cibles sont particulièrement méconnus chez le sorgho.

1.2.4. Objectifs des analyses de réseaux de co-expression de gènes

Les FT qui régulent la mise en place des parois chez le sorgho sont donc en grande majorité inconnus. Les connaissances acquises chez d'autres espèces couplées à une analyse phylogénétique globale permettront d'identifier des FT MYB et NAC potentiellement clés dans la régulation des parois chez le sorgho. L'utilisation d'un réseau de co-expression pour analyser les groupes de gènes dont ces gènes candidats font partie permettra d'accumuler un autre niveau d'information pour leur rôle chez le sorgho. Au-delà de l'obtention d'indices sur le rôle de gènes dont on connaît déjà l'importance chez d'autres espèces, cette approche permet une analyse globale de tous les gènes des familles impliquées et la mise en évidence de nouveaux FT potentiellement régulateurs de la mise en place des parois chez le sorgho. L'analyse des groupes de co-expression peut être réalisée grâce à différentes approches. L'étude de l'enrichissement en « GO-terms » (Gene Ontology, GO) de ces groupes permet d'identifier les grands processus biologiques, et en particulier ceux de mise en place des parois, dans lesquels les gènes qui le composent sont impliqués. L'enrichissement en gènes de biosynthèse des parois des groupes de gènes est également testé dans cet objectif. Les gènes clés des groupes identifiés comme étant impliqués dans la mise en place des parois peuvent être identifiés grâce à des mesures de centralité comme la connectivité. Ce sont des gènes potentiellement cruciaux pour la mise en place des parois, que ce soit dans la régulation ou dans la biosynthèse.

L'utilisation des FT identifiés par les approches sur les données issues de transcriptomes de sorgho mentionnées ci-dessus et par leur rôle chez d'autres espèces peuvent servir d'ancrage solide à l'analyse des autres gènes qui composent leurs groupes de co-expression. En particulier, les gènes dont l'expression est la plus corrélée avec ces FT sont des gènes également potentiellement impliqués dans la mise en place des parois.

Le travail décrit ici a donc pour objectif de renforcer par l'analyse d'un réseau de co-expression de gènes les informations suggérant l'implication de FT des familles NAC et MYB déjà identifiés chez d'autres espèces dans la mise en place des parois chez le sorgho. L'identification de nouveaux FT et de listes de gènes clés est également attendue.

1.3. Génétique d'association

Les approches de phylogénétique et d'analyse de réseaux de co-expression de gènes permettent d'identifier de nombreux gènes candidats dont la fonction est incertaine. Dans le

cadre de ces approches, la comparaison de listes de gènes candidats obtenus de différentes manières permet l'accumulation d'informations qui peuvent permettre de canaliser un faisceau de preuves sur certains gènes candidats. L'utilisation de résultats de génétique d'association est une analyse complémentaire qui permet également d'explorer la diversité génétique des gènes mis en évidence dans l'objectif d'identifier leur potentiel en sélection.

1.3.1. Principe des études de génétique d'association

Les analyses de génétique d'association reposent sur la formalisation mathématique de la relation qui existe entre génotype et phénotype. Le phénotype étant la réalisation du génotype dans un environnement donné, il est nécessaire d'observer le génotype dans différents environnements ou à travers différents individus pour caractériser au mieux la relation qui existe entre génotype et phénotype. Chez les plantes, la possibilité de produire de nombreux descendants issus de croisements contrôlés et de conduire des essais répliqués est très favorable à la mise en place d'études d'association dans l'objectif d'évaluer précisément le génotype en lien avec les caractéristiques d'intérêt (Zhu *et al.*, 2008).

L'analyse de population apparentées (biparentales et multiparentales) permet l'identification de Quantitative Trait Locus (QTL), des zones génétiques de plusieurs milliers voire millions de kilobases associées aux caractères phénotypiques évalués dont la précision est limitée par le nombre de recombinaisons se produisant entre les individus. L'analyse de populations diverses à large base permet, elle, la mise en place de Genome Wide Association Studies (GWAS) dont la résolution est meilleure (Figure II-1). En effet, la grande diversité génétique des populations couplée à une histoire évolutive profonde (i.e. de nombreux événements de recombinaisons) ainsi que l'utilisation d'un grand nombre de marqueurs génétiques permet d'identifier des locis très précis impliqués dans la variabilité des caractères évalués sous la forme notamment de Single Nucleotide Polymorphism (SNP).

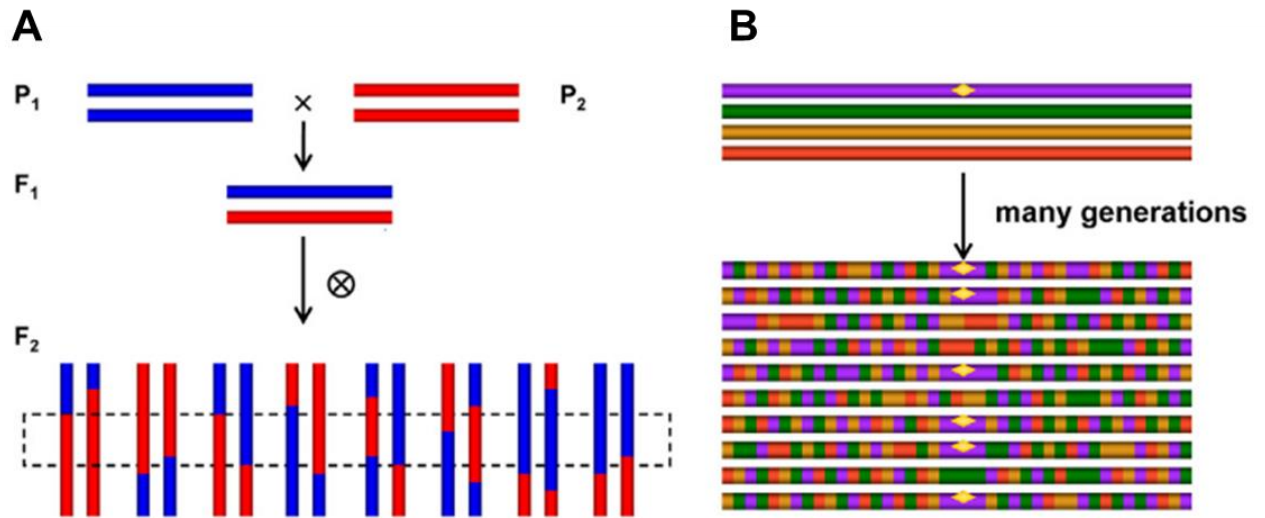


Figure II-1: Comparaison schématique des analyses d'association menées sur des populations biparentales et sur des populations à diversité large (D'après Zhu *et al.*, 2008)

Les identifications de QTL ont été très utiles pour la sélection de caractères d'intérêt agronomiques (Jiang, 2015) mais les analyses GWAS permettent d'identifier des zones plus petites du génome ce qui est favorable à l'identification des gènes responsables des variations de phénotypes évaluées. L'identification des gènes est cependant soumise à la limitation du déséquilibre de liaison qui qualifie la liaison génétique entre deux gènes qui ségrégent non indépendamment l'un par rapport à l'autre. Ce phénomène peut entraîner des intervalles de confiance relativement grands qui ne permettent pas d'isoler des gènes individuels.

1.3.2. Associations détectées chez le sorgho en lien avec la qualité de la biomasse

Chez le sorgho, des analyses d'association Phénotype-Génotype ont été menées pour comprendre les variations phénotypiques des parois secondaires. Par exemple, concernant le contenu en lignine, de nombreux auteurs ont montré (Murray *et al.*, 2008a; Shiringani et Friedt, 2011; Trouche *et al.*, 2014) qu'une grande part de la variation génétique détectée pour ce trait est indépendante des effets des gènes Bmr.

Murray et al (2008a, 2008b) ont été les premiers à analyser l'architecture génétique de la composition de la biomasse et en particulier des carbohydrates de structures et du contenu en lignines dans les tiges et les feuilles en utilisant une population biparentale. Ils ont observé une grande instabilité des QTL détectés à travers les divers environnements et mis en évidence l'indépendance des contrôles génétiques de ces caractères entre les feuilles et les tiges. Enfin, ils ont également rapporté de nombreuses co-localisations entre les QTL détectés pour la composition des parois secondaires et les caractéristiques phénologiques et en composition de la biomasse. Shiringani et Friedt (2011) sont parvenus aux mêmes conclusions avec le même type de population et ont également mis en évidence des colocalisations avec des QTL détectés par Murray *et al.* (2008b, 2008a) Une troisième analyse basée sur une population biparentale interspécifique (*Sorghum bicolor ssp bicolor* x *Sorghum propinquum*) ciblant le potentiel d'hydrolyse des polysaccharides et la cristallinité de la cellulose (Vandenbrink *et al.*, 2013) a également reporté majoritairement des contrôles génétiques indépendant de ces deux traits dans les feuilles et les tiges et ont échoué à détecter des QTL communs à ces caractères. Brenton *et al.* (2016) ont également analysé une population composée d'accessions de sorghos sucrés et de sorghos biomasse dans l'objectif d'identifier des locus contrôlant le contenu en parois de la biomasse (à travers une mesure du NDF) et ont mis en évidence des loci communs avec ceux responsables de la variabilité du contenu en

carbohydrates. Wang et al (2013) ont identifié un gène connu de la régulation des parois secondaires, NST1, sous-jacent à une zone d'association liées au rendement de saccharification chez le sorgho. Un autre gène NAC074 a également été mis en évidence et est associé à la jutosité des tiges (Casto *et al.*, 2018; Xia *et al.*, 2018; Zhang *et al.*, 2018). Les régions génétiques associées à la variabilité génétique de la composition des parois sont donc nombreuses mais n'ont pas, ou peu, permis l'identification de gènes précis liés aux associations détectées.

1.3.3. Objectifs de l'analyse des associations Phénotype-Génotype

Les approches de génomique comparative et d'analyse de réseaux de gènes prévues permettront l'accumulation d'informations sur le rôle des FT NAC et MYB ainsi que sur les gènes de structure qui sont intégrés dans leurs réseaux de régulation. Un renforcement des conclusions relatives à leur implication réelle dans la mise en place des parois secondaires sera permis par l'analyse des gènes communs entre listes de gènes candidats mis en évidence par ces approches et celles de génétique d'association. Ainsi, la comparaison des listes de gènes candidats issues de génétique d'association, les FT et leurs gènes co-régulés ainsi que l'ensemble des gènes présents dans des groupes de co-expression liés aux parois est l'aboutissement de la recherche de gènes candidats impliqués dans l'établissement des parois secondaires chez le sorgho présentée dans cette thèse.

2. Matériels et méthodes

2.1. Analyses phylogénétiques des familles MYB et NAC

Les données publiques des séquences protéiques MYB et NAC d'*Arabidopsis thaliana* ont d'abord servi à produire des signatures protéiques de chaque famille en utilisant les programmes MAAFT (Katoh *et al.*, 2017) et TrimAl (Capella-Gutiérrez *et al.*, 2009) intégrés à la boîte à outils Galaxy (Afgan *et al.*, 2018) (seuls les sites avec moins de 90% d'intervalles vides et au moins 10% du total des sites ont été conservés, les valeurs par défaut des autres paramètres ont été utilisées). La deuxième étape a consisté à sélectionner des seuils de similarité pour agréger les protéines des différentes familles chez les différentes espèces (E-value= 10^{-22} pour la famille des MYB et E-value= 10^{-30} pour la famille des NAC). Ces valeurs ont été déduites des valeurs de E-value observées pour récupérer les protéines connues d'*Arabidopsis* pour ces deux familles. Ces signatures « *Arabidopsis* » ont ensuite été utilisées pour récupérer les protéines homologues issues des protéomes de 8 espèces additionnelles en

utilisant l'approche HMM (Hidden Markov Model). Cinq protéomes de monocotylédones (le maïs: *Zea mays*, le riz: *Oryza sativa* susp *japonica*, le sorgho: *Sorghum bicolor*, brachypodium: *Brachypodium distachyon*, et le millet des oiseaux: *Setaria italica*) et 4 protéomes de dicotylédones (en plus d'Arabidopsis: *Arabidopsis thaliana*, luzerne: *Medicago truncatula*, eucalyptus: *Eucalyptus grandis*, et peuplier: *Populus trichocarpa*) ont été utilisés. Ces espèces ont été choisies car elles sont des modèles d'études des parois secondaires.

Ensuite, les alignements des protéines complètes ont été générés avec MAAFT et les séquences ont été nettoyées avec TrimAl (option gt=0,9 et option cons=10 à nouveau). Les alignements ont été triés en gardant le plus long variant d'épissage de chaque gène. Un arbre phylogénétique par famille de FT a été généré avec la méthode du maximum de vraisemblance implémentée dans PhyML (Guindon *et al.*, 2010; Zhou *et al.*, 2018). L'outil d'enracinement d'arbre Rap-Green (Dufayard *et al.*, 2005) a été utilisé pour réconcilier les arbres obtenus avec l'arbre des espèces afin de pouvoir identifier les duplications de gènes et optimiser l'inférence de fonction de gènes. Les arbres phylogénétiques sont disponibles en ligne à l'adresse <https://phylogeny.southgreen.fr/Hennetetal/>

2.2. Réseaux de co-expression

2.2.1. Evolution de l'expression des gènes pendant le développement de la tige

i. Conditions expérimentales au champ et stratégie d'échantillonnage

L'évolution du transcriptome d'entre-nœuds en développement a été analysée au sein de trois expériences au champ en 2013, 2014 et 2015. Durant ces trois expérimentations, « Biomass 140 », un hybride commercial développé par Eurosorgho pour sa forte production de biomasse et sa floraison tardive, a été étudié en conditions contrastées de disponibilité en eau sur la plateforme de phénotypage DIAPHEN à Mauguio (Delalande *et al.* 2015; 43°36'43''N, 3°58'2'' E) pendant l'été (date de semis les 22, 23 et 13 mai respectivement pour 2013, 2014 et 2015).

Les conditions d'approvisionnement en eau correspondent aux traitements « Well-Watered » (WW – Bon apport en eau) et « Water Deficit » (WD – Déficit en eau). WD a consisté en un mois de sécheresse progressive qui a commencé au stade moyen de 11 feuilles ligulées développées sur la tige principale. Ce stade a été choisi car il correspond au début de l'élongation des entre-nœuds (Gutjahr *et al.*, 2013). Les expérimentations correspondent à des

designs expérimentaux en blocs aléatoires avec 4 et 3 réplicas respectivement en 2013 et 2014-2015. Les parcelles individuelles étaient composées de lignes 7m de long espacées de 80 cm avec 8 rangées par parcelle en 2013 et 4 en 2014-2015. Dix-huit graines ont été semées au mètre linéaire.

Des stratégies différentes d'échantillonnage ont été utilisées pour les trois années d'expérimentation. En 2013, deux niveaux d'entre-nœuds correspondant aux feuilles ligulées 12 et 16, ont été prélevés au cours de 5 stades de développement (correspondant aux stades D1 à D5) pour les deux régimes d'irrigation (à l'exception de l'entre-nœud du haut pour le premier stade car il n'était pas encore présent). En 2014, 6 niveaux d'entre-nœuds ont été prélevés au cours de 2 stades de développement dans les deux régimes d'irrigation. Les entre-nœuds correspondant à la 13^{ème} et 15^{ème} feuille ligulée ont été prélevés aux cours des 2 stades. Les entre-nœuds des 14^{ème} et 16^{ème} feuilles ligulées furent prélevés uniquement au premier stade et ceux des 17^{ème} et 19^{ème} feuilles ligulées uniquement au dernier stade. En 2015, 4 niveaux d'entre-nœuds ont été prélevés au cours de 2 stades de développement dans les deux régimes d'irrigation. Dans ce cas, les entre-nœuds ont été prélevés selon leur position par rapport à la dernière feuille ligulée de la tige. A la première date de prélèvement (23 juillet 2015), les 4 premiers entre-nœuds les plus hauts ont été prélevés. En moyenne, ils correspondaient aux rangs de 14^{ème}, 13^{ème}, 12^{ème} et 11^{ème} entre-nœud. Pour la deuxième date de prélèvement (24/09/2015), les 4 entre-nœuds prélevés correspondaient à l'entre-nœud de la feuille drapeau (Flag Leaf=FL, ce qui correspond en moyenne à l'entre-nœud 23), à FL-2 (entre-nœud 21), FL-4 (entre-nœud 19) et FL-6 (entre-nœud 17). Chaque échantillon d'entre-nœud correspond à un groupe de 3 ou 4 entre-nœuds individuels issus de plantes indépendantes pour 2014-2015 et 2013 respectivement. Des échantillons d'entre-nœud dupliqués ont été prélevés pour chaque niveau d'entre-nœud, stade et réplicas biologique puis soit directement congelés dans de l'azote liquide et conservés à -80°C après la récolte pour des analyses transcriptomiques, soit séchés à 60°C pendant 72h pour des analyses spectrales de leur composition.

ii. Isolement des ARN totaux et construction des banques

L'isolement des ARN totaux a été effectué en suivant la procédure décrite dans Chomzynski et Sacchi (1987). Les entre-nœuds congelés (-80°C) ont été broyés en une fine poudre en utilisant le moulin IKA A11. Un millilitre de TRIZOL[®] Reagent a été utilisé pour 100 mg de poudre. Les RIN (RNA IntegrityNumber) et les quantifications d'ARN totaux ont été mesurées en utilisant la puce Agilent 2100 Bioanalyzer Nano 6000. Un total de 350 banques

RNAseq (158 en 2013, 96 en 2014, 96 en 2015) ont été préparées en suivant le protocole Illumina avec le kit TruSeq RNA Library Prep Kit (Illumina, USA), en utilisant 1µg des RNA totaux. Les banques indexées ont été groupées par 24 et ont subi un séquençage en single-end (1x150pb) sur un Illumina HiSeq2500 (Plateforme Genotoul à Toulouse, France). Chaque groupe de 24 banques a été séquençé en parallèle sur deux lignes permettant d'obtenir un réplicas technique (14 lignes au total). Les séquences brutes sont disponibles sur SRA (short read archive) (<https://dataview.ncbi.nlm.nih.gov/object/PRJNA560153?reviewer=8kt9rf7e6hfi38a8asq9kgnuvb>).

iii. Contrôle qualité, alignement et comptage du nombre de séquences obtenues par transcrit

FastQC (Andrews, 2010) a été utilisé pour vérifier la qualité des séquences brutes (i.e. détection des adaptateurs, qualité moyenne de la séquence complète (score phred), le taux de GC, la duplication des séquences, la contamination des PCR). Les séquences obtenues ont été nettoyées en utilisant cutAdapt (Martin, 2011) pour éliminer l'extrémité des séquences de mauvaise qualité (Score Q en dessous de 20) et pour garder uniquement les séquences de qualité moyenne supérieure à 30 et d'une taille minimale de 35 paires de bases. Pour ces étapes, les programmes décrits par Sarah *et al.* (2017) et disponibles sur la plateforme Southgreen ont été utilisés : `arcad_hts_0_fastqc_in_chains.pl`, `arcad_hts_1_cutadapt_in_chain.pl`, `arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl`

Le niveau d'expression des transcrits a été estimé avec le script « New-Tuxedo » (Pertea *et al.*, 2016). Tout d'abord, pour chaque échantillon d'ARN, les séquences ont été cartographiées sur la séquence de référence du sorgho Sbicolor_313_v3.1 en utilisant Hisat2 (Kim *et al.*, 2015). Les transcrits ont été assemblés et quantifiés avec Stringtie (Pertea *et al.*, 2016) en utilisant le fichier d'annotation de référence (au format GFF3) qui permet de guider le processus d'assemblage. La sortie consiste en des transcrits de référence (i.e. déjà disponibles sur le génome de référence) exprimés ainsi que tous les nouveaux transcrits assemblés. Gffcompare (ccb.jhu.edu/software/stringtie/gff.shtml#gffcompare) a été utilisé pour comparer les transcrits avec l'annotation de référence (prédictions de gènes et de transcrits) et pour identifier les nouveaux gènes ou transcrits. Les fichiers assemblés et enrichis de ces annotations ont été utilisés pour estimer leur abondance avec Stringtie.

2.2.2. Analyse des réseaux de co-expression de gènes

i. Identification du réseau de co-expression de gènes

Seuls les gènes présentant des niveaux d'expression supérieurs à 3 transcrits dans au moins 20 échantillons d'entre-nœud (sur les 350 échantillons) ont été considérés pour les analyses de co-expression. Les réseaux de co-expression de gènes ont été construits en utilisant le package R WGCNA (Langfelder et Horvath, 2008, 2012) en utilisant les données normalisées et filtrées. La normalisation a été faite avec le package EdgeR (Robinson *et al.*, 2010). Un total de 20.294 gènes a été retenu pour l'analyse des réseaux de co-expression de gènes. Les effets bloc et années ont été estimés en utilisant la fonction « removeBatchEffect » du package R « Limma » (Ritchie *et al.*, 2015). Aucun effet de ces facteurs n'a été détecté et aucun échantillon aberrant n'a été identifié. Le réseau a été construit en utilisant le paramètre networkType « signed » qui permet de prendre en considération le sens de la variation d'expression en groupant les gènes présentant le même sens de variation. Ce paramètre est recommandé pour obtenir des modules de gènes qui ont un sens biologique (van Dam *et al.*, 2018). En se basant sur les courbes de la connectivité moyenne et de l'index scale-free topology (Annexe SII-1A), un pouvoir de 12, qui est classiquement utilisé pour ce type de données, a été utilisé pour nos analyses. Cela montre que nos données sont en accord avec les modèles d'approximation de scale-free topology, ce qui est requis pour réaliser une analyse avec WGCNA. Cela indique qu'aucun biais fort n'existe dans nos données. Les relations entre chaque paire de gènes ont été explorées en utilisant un coefficient de corrélation de Pearson. Seuls les gènes avec de fortes corrélations d'expression avec d'autres gènes ont été considérés. Un coefficient de proximité (qui est une corrélation à la puissance du pouvoir déterminé, ce qui permet d'amplifier les disparités entre les corrélations faibles et fortes) au seuil de 0.1 a été utilisé ce qui correspond à une corrélation d'expression minimum de 0.82 entre deux gènes. Autrement, tous les paramètres par défaut de WGCNA ont été utilisés. Les modules sont nommés d'après des couleurs choisies aléatoirement.

ii. Analyse des enrichissements des réseaux de gènes

Pour une analyse globale, les modules obtenus avec WGCNA ont été caractérisés en se basant sur leur enrichissement en GO-terms (Gene Ontology, GO) en utilisant le package R TopGO (Alexa et Rahnenfuhrer, 2019). Les paramètres par défaut et la liste de GO du sorgho issue de la base de données PLAZA monocots 4.0 ont été utilisés. Une p-value de 0.05 a été considérée pour définir les GO-terms significativement enrichis.

Nous avons regroupé les GO-terms relatifs aux processus Biologiques (Biological Process) en groupes de mots clés afin de pouvoir les traiter plus facilement (Annexe SII-2). Par exemple, le mot clé « SCW » (Secondary Cell Wall, parois secondaires) regroupe 4 Go-terms différents: negative/positive “regulation of secondary cell wall biogenesis”, “regulation of secondary cell wall biogenesis” et “plant-type secondary cell biogenesis”.

Une liste de gènes de structure liés aux parois a été compilée pour les besoins de cette étude à partir de listes de gènes issues de différentes publications traitant du développement de l'entre-nœud chez le sorgho (Kebrom *et al.*, 2017; McKinley *et al.*, 2016; Petti *et al.*, 2015; Rai *et al.*, 2016). En plus de l'analyse de l'enrichissement en GO-terms, des tests exacts de Fisher ont été réalisés pour chacun des modules pour détecter un enrichissement en ces gènes de structure liés aux parois.

De plus, les gènes sorgho MYB et NAC (SbMYB et SbNAC) identifiés dans les analyses phylogénétiques ont été utilisés comme gènes d'accroche pour caractériser les réseaux de gènes identifiés. Dans cet objectif, pour chacun des gènes SbMYb et SbNAC, les 15 gènes dont l'expression leur est la plus corrélée ont été extraits. Puis un deuxième niveau de gènes co-exprimés ont été extraits en récupérant à nouveau les 15 gènes les plus co-exprimés avec le premier niveau de gènes. Les réseaux de régulation étant un ensemble de cascades de régulation, cette approche tente d'avoir un aperçu de la diversité des gènes impliqués dans le réseau au-delà du premier niveau de régulation. Les gènes de ces « sous-réseaux » ont été caractérisés en analysant leurs enrichissements en GO-terms et en gènes liés aux parois présentés précédemment.

iii. Identification des « hub-genes »

La connectivité intramodule (Intramodular connectivity) de chaque gène a été calculée en utilisant la fonction WGCNA “intramodularConnectivty.fromExpr”. Plus haute la connectivité est, plus le gène a une importance centrale dans le réseau. Ces gènes, appelés « hub-genes », sont donc supposés jouer des rôles biologiques cruciaux. Nous avons analysé les 10 meilleurs gènes présentant la plus forte connectivité de chaque module afin d'avoir un aperçu précis des gènes cruciaux de ces modules ainsi que les gènes faisant partie des 10% ayant le meilleur score pour chaque module. Cette dernière analyse, moins facile à appréhender que celle des 10 gènes avec les plus fortes connectivités, permet d'avoir une idée réaliste des gènes centraux des réseaux puisque la position dans le classement compte moins

que d'être globalement dans le haut du classement (FAQ WGCNA : <https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/faq.html>).

Nous avons comparé ces deux listes de gènes (10 meilleurs et 10% meilleurs) pour chaque module en considérant ou non les seuils de corrélation et d'expression (groupes d'interactions « forts » ou totaux). Le nombre de gènes connectés par les réseaux entre en compte dans le calcul du score de la connectivité, il est donc possible que la sélection d'un sous-ensemble de gènes du module (seuils d'expression et de corrélation) ait un impact sur celui-ci et ne reflète pas correctement la connectivité des gènes.

iv. Caractérisation biochimique des entre-nœuds

Les quantifications des contenus en lignines, cellulose et hémicelluloses ont été effectuées grâce à la méthode de référence Van Soest (1991) pour laquelle des calibrations en Spectrométrie proche Infra Rouge ont été développées. La méthode Van Soest fournit des estimations des concentrations en fibres totales (NDF exprimé en pourcentage de la Matière Sèche, % MS, de la ligno-cellulose (ADF en % MS) et des lignines (ADL en % MS). Les mêmes niveaux d'entre-nœuds utilisés pour les analyses transcriptomiques ont été échantillonnés et séchés pendant 72°C à 60°C. Les échantillons séchés ont été broyés avec une maille de tamis 1 mm et les spectres en proche infra-rouge ont été acquis grâce à un spectromètre NIR système 6500 (FOSS NirSystem, Laurel, MD, United States). Les modèles de calibration pour les différents caractères ont été développés grâce à un ensemble de 700 mesures biochimiques pour chacun d'eux. NDF, ADF et ADL ont été utilisés pour calculer la teneur en hémicellulose (NDF-ADF) et la teneur en cellulose (ADF-ADL).

2.2.3. Analyse de génétique d'association

i. Panels de GWAS

Deux panels mondiaux de diversité ont été analysés dans le cadre de ce travail: la core-collection du CIRAD (CC) contenant 210 accessions développée par Deu *et al.* (2006) et le Set de Référence du Generation Challenge Program (RSGCP) contenant 384 accessions décrits par Billot *et al.* (2013). Ces deux panels présentent 40% d'intersection, tous deux couvrant une grande diversité génétique, toutes les races de sorgho cultivées et la plupart des aires géographiques d'origines des sorghos cultivés. En sus, le panel RSGCP inclut quelques accessions de *S. bicolor* sauvages de la sous espèce *verticilliflorum*. Les accessions ont été

considérées en majorité homozygotes malgré une faible hétérozygotie résiduelle à quelques loci pour certaines accessions sauvages.

ii. Dispositifs expérimentaux des essais de terrain

Les panels ont été évalués pour la composition et la production de biomasse en 2008 et 2009 sur trois sites: en France à Montpellier Lavellette station expérimentale du GEVES (43°35'N, 3°52'E), au Mali dans la station de l'Institut d'Economie Rurale (IER) de Sotuba (12°39' N, 7° 56' W) et au Mali à la station ICRISAT de Samanko (12° 32' N; 8° 7' W).

Les deux essais menés en 2008 et 2009 à la station de Lavalette (Montpellier, France) ont été décrits par Trouche *et al.* (2014). Ces essais étaient structurés en blocs augmentés avec 3 témoins en 2008 et 4 en 2009 pour chacun des blocs. Pour ces deux essais seules les accessions non-photopériodiques ont été utilisées car les autres n'étaient pas capables de fleurir dans les conditions de lumière et de température de l'essai.

L'essai à Samanko au Mali a été mené en blocs augmentés et a inclu 10 blocs et 4 variétés témoins qui ont été répliquées dans chacun des blocs. Cet essai a été géré pendant la contre-saison avec une irrigation artificielle. Les deux essais menés à Sotuba au Mali étaient des designs alpha-plans avec deux réplifications.

iii. Phénotypage

Les caractères présentés dans le Tableau II-1, regroupés par types généraux de caractères (méta-trait) ont été prédits pour ces différents essais par spectrométrie proche infrarouge (NIRS).

Seules les caractéristiques phénotypiques liées à la tige ont été retenues pour les analyses présentées ici. Ainsi, les méta-traits « Durée du cycle » et « Matière sèche » n'ont pas été retenus et les méta-traits « Digestibilité animale », « Paroi végétale », et « Sucres et jus » ont été considérés.

Tableau II-1 : Caractères mesurés par spectrométrie proche infrarouge (NIRS) et regroupés par type général de caractère (méta-trait)

Meta-trait	Abréviation du caractère	Description
Production de matière sèche	DSW	Biomasse de la tige principale (g)
	PHT	Hauteur de la plante (cm)
	DMW	Biomasse de la plante (g)
	DM	Teneur en matière sèche de la tige principale (%)
	MM	Cendre (% DM)
	CP	Protéines brutes (% DM)
Parois végétales	NDF	Neutral Detergent Fiber (% DM)
	ADL	Stem Acid Detergent Lignin (% DM)
	ADF	Stem Acid Detergent Fiber (% DM)
	ADL/NDF	Teneur en lignine de la tige (% NDF)
	CBW	Cellulose de Weende de la tige (% DM)
	CBW/NDF	Teneur en cellulose de Weende de la tige (% NDF)
	CVS	Cellulose de Van Soest de la tige (% DM)
	CVS/NDF	Teneur en cellulose de la tige (% NDF)
	HEMI	Hémicellulose de la tige (% DM)
	HEMI/NDF	Teneur en hémicellulose de la tige (% NDF)
Digestibilité animale	SMO	Dégradation <i>in vitro</i> de la matière organique de la tige
	IVNDFD	Digestibilité <i>in vitro</i> du NDF de la tige (% NDF)
Production de sucres et de jus	BRIX	Brix (%)
	BRIX30	Brix ajusté à 30% DM dans les tiges (%)
	VJUICE	Volume de jus de la tige (ml)
	JUICE	Masse du jus de la tige (gr)
Durée du cycle	DTF	Jours jusqu'à la floraison (j)
	TTF	Temps thermique jusqu'à la floraison (dj)
	DTH	Jours jusqu'à l'apparition du panicule (j)
	TTH	Thermal Time to Heading (dj)
	DFM	Jours de la floraison à la maturité (j)
	TTFM	Temps thermique de la floraison à la maturité (dj)
	DTM	Jours à la maturité (j)
	TTDTM	Temps thermique jusqu'à la maturité (dj)

iv. *Génotypage et attribution des SNP*

Le génotypage a été obtenu en utilisant la technologie de Génotypage par Séquençage (Genotyping-By-Sequencing, GBS : Elshire *et al.*, 2011). La matrice génotypique disponible contient 191 177 SNP associés aux 394 accessions phénotypées. Ces données ont été générées par Ingrid Vilmus (Publication en préparation).

v. *Analyse des données phénotypiques*

Les héritabilités ont été estimées grâce à un modèle sans SNP et d'autres covariables fixées (i.e. les groupes de structure issus de l'analyse dapl) en utilisant le package R « heritability » (Kruijer *et al.*, 2015). Les moyennes génotypiques ajustées ont été utilisées pour tous les essais et données individuelles quand elles étaient disponibles (Mali GCP 2008 et 2009 et France/Montpellier 2008 et 2009).

vi. *Analyse d'association sur le génome entier*

Les analyses d'association sur le génome entier (Genome-Wide Association, GWAS) ont été conduites avec un modèle mixte considérant la structure et l'apparentement en utilisant le modèle Linear Mixed Model (LMM) du logiciel GEMMA (Zhou et Stephens, 2012) pour chaque essai séparément ou en utilisant le blup génotypique pour les essais répliqués au cours des années 2008 et 2009 en France et au Mali. Comme le nombre d'individus était différent sur chaque combinaison d'essai-phénotype, les SNP ont été filtrés avant chaque analyse GWAS en écartant les SNP monomorphiques (causés par la suppression d'individus sans phénotype) et avec une fréquence des allèles mineurs inférieure à 5%. De la même manière, un groupe de structure était enlevé si aucun des individus considérés n'avait des probabilités d'appartenance à ce groupe et la matrice d'apparentement a été réduite pour contenir seulement les individus phénotypés pour la variable considérée dans l'essai.

Les seuils de significativité ont été choisis en utilisant les résultats de l'analyse totale corrigés au taux de fausses découvertes (False Discovery Rate, FDR) de 5%. Comme cette méthode est sensible au type de données, elle ne donne pas un seuil unique mais pour simplifier les interprétations un seuil fixe de significativité de p-value de 10^{-5} a été choisi. Il sera appelé le LOP-5 en utilisant l'abréviation LOP pour la transformation classique $-\log_{10}(\text{p-value})$. A cause de la présence de nombreux essais avec des caractères en commun, nous avons également considéré les associations au LOP-3.

vii. *Identification des zones d'intérêt « hotspot » liées à la composition de la biomasse.*

Les SNP identifiés au moins deux fois au cours de deux essais indépendants ont été sélectionnés à condition que les détections soient effectuées au moins à LOP-5 et l'autre au moins à LOP-3. Les SNP retenus grâce à l'analyse de plusieurs essais (blup) ont été sélectionnés pour une détection seule inférieure aux essais indépendants, à LOP-3. Les SNP détectés ainsi et situés à moins de 5kb les uns des autres ont été assemblés. L'intervalle dans lequel ils sont détectés a été élargi à 20kb de part et d'autres des SNP présents aux extrémités de l'intervalle pour former les zones d'intérêts. La taille de 20kb a été définie de façon conservative sur la base du déséquilibre de liaison chez le sorgho (r^2 de 0.1 à environ 15 kb). Une zone d'intérêt peut être composée d'un seul SNP si sa détection satisfait aux critères mentionnés ci-dessus. La détection des gènes sous-jacents aux zones d'intérêt a été réalisée sur les données de la version v3.1 du génome du sorgho disponibles sur Phytozome v12.1.

3. Résultats

3.1. Phylogénie des familles NAC et MYB : apport pour l'identification des FT régulant la mise en place des parois secondaires chez le sorgho

3.1.1. Phylogénie des Facteurs de Transcription NAC

Les FT NAC issus de 5 espèces de monocotylédones (sorgho, maïs, riz, mil, brachypodium) et 4 dicotylédones (Arabidopsis, peuplier, eucalyptus, luzerne) ont été récupérés à partir de bases de données protéomiques des génomes complets. Cent douze gènes NAC ont été identifiés chez Arabidopsis, 168 chez le peuplier, 163 chez eucalyptus, 70 chez la luzerne, 136 chez le maïs, 132 chez le mil, 129 chez le riz, 138 chez brachypodium et 122 chez le sorgho. Ces nombres sont relativement similaires à ceux qui ont déjà été reportés dans la littérature tout en étant plutôt dans la partie basse de la distribution. Précédemment, 105 (Ooka *et al.*, 2003) et 117 (Nuruzzaman *et al.*, 2010) gènes NAC ont été identifiés chez Arabidopsis, 288 (Pereira-Santana *et al.*, 2015) et 163 (Hu *et al.*, 2010) chez le peuplier, 189 chez l'eucalyptus (Hussey *et al.*, 2015), 97 chez la luzerne (Ling *et al.*, 2017), 124 (Fan *et al.*, 2014) et 152 (Shiriga *et al.*, 2014) chez le maïs, 147 chez le mil (Puranik *et al.*, 2013), 151 (Nuruzzaman *et al.*, 2010) et 140 (Fang *et al.*, 2008) chez le riz, 107 (You *et al.*, 2015) et 118

(Zhu *et al.*, 2015) chez brachypodium et enfin 131 (Sanjari *et al.*, 2019) et 145 (Kadier *et al.*, 2017) gènes NAC chez le sorgho.

Nous avons utilisé une étude récente de la famille des NAC chez le sorgho (Sanjari *et al.*, 2019) afin de nous aider à caractériser l'arbre phylogénétique que nous avons obtenu en compilant les séquences des gènes NAC de ces 9 espèces. Nous avons retrouvé 122 des 131 gènes NAC reportés dans la publication de Sanjari *et al.* (2019) et 8 des 15 clades de gènes sorgho identifiés dans cette dernière sont inchangés dans notre arbre (Figure II-2A). L'ancien clade N a été fusionné à l'ancien clade O pour former le nouveau clade O. Le clade K a été divisé en deux nouveaux clades, S et K. Le clade L a été lui aussi divisé en deux sous-clades 1 et 2. Le nouveau clade F rassemble la plupart des gènes de l'ancien clade F ainsi que quelques gènes de clades différents et tous les gènes des anciens clades I et H qui n'existent plus. Parmi les 9 gènes reportés dans l'étude de Sanjari *et al.* (2019) et non identifiés au cours de notre travail, 7 sont des gènes de l'ancien clade F. En dehors de ces gènes, seuls 4 gènes n'ont pas une position qui correspond aux clades définis par Sanjari *et al.* (2019) et ont été alloués à d'autres clades dans notre classification. De surcroît, un nouveau clade P a été créé car il ne contient que des gènes de monocotylédones et aucun gène sorgho permettant de le rattacher à l'étude de Sanjari *et al.* (2019).

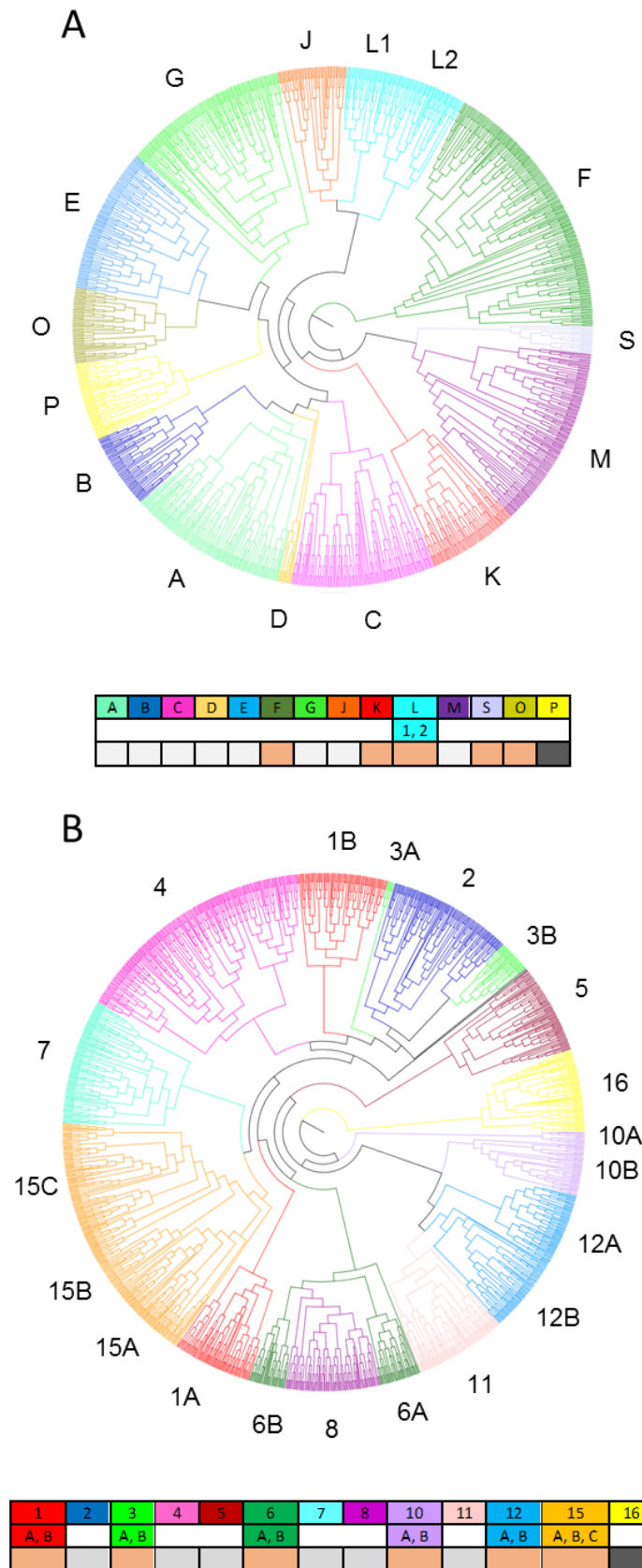


Figure II-2: Arbres phylogénétiques simplifiés des NAC (A) et des MYB (B) construits avec les séquences protéiques de 5 monocotylédones et 4 dicotylédones. Les couleurs correspondent aux différents clades (des lettres pour les NAC et des chiffres pour les MYB, en légende). Les sous-clades sont désignés avec des chiffres pour les NAC et des lettres pour les MYB. Les comparaisons menées avec les travaux de Sanjari *et al.* (2019) pour les NAC et Li *et al.* (2016) pour les MYB sont fournies sous les identifiants de clades. Les clades conservés sont en gris clair, les nouveaux clades sont en gris foncé et les clades avec réallocation de gènes sont en orange.

Un enrichissement des clades B et F en gènes de monocotylédones a été observé. Le clade F, qui est le plus gros clade défini avec 193 séquences, contient 135 (69%) gènes de monocotylédones et 58 (31%) gènes de dicotylédones. Aucune information fonctionnelle chez *Arabidopsis* n'est disponible pour ce clade. Le clade B contient 81% de gènes monocotylédones (44 d'un total de 54). D'après la fonction reportée de AtNAC1 qui appartient à ce clade, ces gènes pourraient être impliqués dans le transport de l'auxine (Xie *et al.*, 2000).

Parmi les 13 gènes *Arabidopsis* pour lesquels il existe des évidences fonctionnelles d'un rôle dans la régulation des parois, 9 appartiennent au clade C, deux au clade G (SND2, SND3) et un au clade K (VNI2).

Globalement, en se basant sur les lignages monocotylédones et dicotylédones détectés pour ces gènes NAC régulateurs de la mise en place parois chez *Arabidopsis*, aucune relation d'orthologie entre *Arabidopsis* et le sorgho n'implique un orthologue unique (Tableau II-2). De plus, à l'exception de VNI2 dont il existe de nombreux homologues chez les monocotylédones (6 homologues chez le sorgho et jusqu'à 13 chez *brachypodium*) ce qui indique une grande diversification, une bonne homogénéité entre les nombres de gènes existe entre *Arabidopsis* et le sorgho. Par exemple, les gènes NST1, SND1 et NST2 ont deux homologues chez le sorgho (Tableau II-2, Figure II-3).

Tableau II-2: Régulateurs MYB et NAC majeurs des parois secondaires chez Arabidopsis et leurs homologues chez le sorgho ainsi que leurs homologues validés fonctionnellement chez d'autres espèces. La méthode de validation et la publication sont également rapportées

TF	Arabidopsis SCW gene			Clade	Homologs			
	Gene ID	Name synonym	Publication		Gene ID	Name synonym	Validation	Publication
NAC	AT2G46770; AT1G32770; AT3G61910	NST1; NST3/SN D1; NST2	Zhong <i>et al.</i> , 2006, 2007, Mitsuda <i>et al.</i> , 2007	C	Zm00001d045463 LOC_Os06g04090	ZmSWN1 OsSWN1	Complète le mutant <i>snd1/nst1</i> chez Arabidopsis et le surexprimeur active des FT régulateurs des parois secondaires	Zhong <i>et al.</i> , 2011
	AT4G28500; AT1G28470	SND2; SND3	Hussey <i>et al.</i> , 2011	G	Sobic.007G018100 Sobic.010G022400	SbSWN2/SbNSTa SbSWN1/SbNSTb		
	AT1G71930	VND7	Yamaguchi <i>et al.</i> , 2010	C	LOC_Os08g01330 JN634079	OsSWN3 ZmSWN3	Complète le mutant <i>snd1/nst1</i> chez Arabidopsis et le surexprimeur active des FT régulateurs des parois secondaires	Zhong <i>et al.</i> , 2011
	AT1G12260; AT1G62700; AT5G62380	VND4; VND5; VND6	Ohashi-Ito <i>et al.</i> , 2010	C	Sobic.007G003000 Sobic.006G279400	SbSWN3/SbVND7a SbSWN2/SbVND7b		
	AT1G12260; AT1G62700; AT5G62380	VND4; VND5; VND6	Ohashi-Ito <i>et al.</i> , 2010	C	Sobic.010G002900 Sobic.006G160900 Sobic.004G302400 JN634082 JN634083 LOC_Os06g01480	SbSWN5/SbVND4-6a SbSWN6/SbVND4-6c SbSWN7/SbVND4-6b ZmSWN6 ZmSWN7 OsSWN7	Complète le mutant <i>snd1/nst1</i> chez Arabidopsis et le surexprimeur active des FT régulateurs des parois secondaires	Zhong <i>et al.</i> , 2011
	AT2G18060; AT4G36160; AT5G66300	VND1; VND2; VND3	Zhou <i>et al.</i> , 2014	C	Sobic.001G316800 Sobic.001G522700	SbVND1-3a SbVND1-3b		

MYB	AT4G29230	NAC075	Endo <i>et al.</i> , 2015	G	Sobic.003G035100 Sobic.006G004100 Sobic.006G003800 Sobic.009G071600 Sobic.009G071500	SbNAC75a SbNAC75b SbNAC75c SbNAC75d SbNAC75e	
	AT5G13180	VNI2	Yamaguchi <i>et al.</i> , 2010	K	Sobic.002G259600 Sobic.007G190333 Sobic.003G409800 Sobic.003G423200 Sobic.006G141900 Sobic.008G094700 Sobic.005G056300	SbVNI2g SbVNI2f SbVNI2e SbVNI2d SbVNI2c SbVNI2b SbVNI2a	
	AT3G08500; AT5G12870	MYB83; MYB46	Zhong <i>et al.</i> , 2007, McCarthy <i>et al.</i> , 2009, Zhong and Ye, 2012	4	Sobic.008G112200 Zm00001d023931 LOC_Os12g33070 Potri.001G267300 Potri.009G061500 Potri.001G258700 Potri.009G053900 Eucgr.G03385	SbMYB83 ZmMYB46 OsMYB46 PtrMYB3 PtrMYB20 PtrMYB21 PtrMYB2 EgMYB2	La surexpression chez Zhong <i>et al.</i> , Arabidopsis active le 2011 programme de biosynthèse des parois La surexpression mène au dépôt McCarthy <i>et al.</i> , ectopique de parois secondaires 2010 Zhong <i>et al.</i> , 2013 La surexpression chez le tabac Goicoechea <i>et al.</i> , 2005 augmente l'épaisseur des parois secondaires
	AT1G09540; AT4G01680	MYB61; MYB55	Newman <i>et al.</i> , 2004, Romano <i>et al.</i> , 2012	4	Sobic.009G036500 Sobic.003G136600 LOC_Os01g18240	SbMYB61a SbMYB61b OsMYB55/61	Les plantes RNAi plants ont un Hirano <i>et al.</i> , phénotype anormal 2013b
	AT1G63910	MYB103	Zhong <i>et al.</i> , 2008	4	Sobic.007G039100 LOC_Os08g05520	SbMYB103 OsMYB103	Les plantes RNAi plants ont un Hirano <i>et al.</i> , phénotype anormal 2013b

AT5G16600; MYB43; Zhong <i>et al.</i> , 2008 AT1G66230 MYB20	4	Sobic.007G132600 SbMYB20/43a Sobic.002G196100 SbMYB20/43b Sobic.002G196000 SbMYB20/43c Sobic.004G248700 SbMYB20/43d Sobic.010G106601 SbMYB20/43e	
AT1G16490; MYB58; Zhong <i>et al.</i> , 2008, AT1G79180 MYB63 Zhou <i>et al.</i> , 2009	1B	Sobic.004G273800 SbMYB60 Sobic.006G199800 SbMYB60b	Scully <i>et al.</i> , 2016a
AT1G17950; MYB52; Zhong <i>et al.</i> , 2008, AT1G73410 MYB54 Zhong and Ye, 2012	12B	Sobic.001G110900 SbMYB52/54	
AT4G33450 MYB69 Zhong <i>et al.</i> , 2008	12B	Sobic.008G063400 SbMYB69a Sobic.005G104800 SbMYB69b	
AT4G12350; MYB42, Zhong <i>et al.</i> , 2008 AT4G22680 MYB85	4	Sobic.002G275500 SbMYB42a Sobic.007G178300 SbMYB42b LOC_Os09g36250 OsMYB42/85 Zm00001d032032 ZmMYB167	Les plantes RNAi plants ont un phénotype anormal La surexpression chez le maïs entraîne une augmentation de la quantité de lignine Hirano <i>et al.</i> , 2013b Bhatia <i>et al.</i> , 2019
AT4G38620; MYB4; Preston <i>et al.</i> , 2004, AT2G16720; MYB7; Mitsuda <i>et al.</i> , 2007, AT4G34990 MYB32 Ko <i>et al.</i> , 2009	15A	Sobic.007G177100 SbMYB4b Sobic.002G279100 SbMYB4a Eucgr.G01774 EgMYB1 Zm00001d053220 ZmMYB42 Zm00001d006236 ZmMYB31	La surexpression chez Arabidopsis réduit l'épaississement des parois secondaires La surexpression réduit le contenu en lignine et inhibe l'enzyme COMT Legay <i>et al.</i> , 2010 Fornalé <i>et al.</i> , 2006

3.1.2. Phylogénie des Facteurs de Transcription MYB

Tout comme pour les gènes NAC, les séquences de 9 espèces ont été récupérées de bases de données comportant leurs génomes complets. Cent quarante-quatre gènes ont été identifiés chez *Arabidopsis*, 210 chez le peuplier, 171 chez l'eucalyptus, 102 chez la Luzerne, 180 chez le maïs, 135 chez le mil, 126 chez le riz, 125 chez le brachypodium et 135 chez le sorgho. Ces effectifs de gènes correspondent globalement à ce qui a déjà été décrit. En effet, précédemment, 197 (Katiyar *et al.*, 2012), 198 (Yanhui *et al.*, 2006) et 133 (Li *et al.*, 2016) gènes MYB ont été identifiés chez *Arabidopsis*, 191 chez le peuplier (Wilkins *et al.*, 2009), 141 chez l'eucalyptus (Soler *et al.*, 2015), 155 chez la luzerne (Zhong et Ye, 2015a), 157 chez le maïs (Du *et al.*, 2012), 209 chez le mil (Muthamilarasan *et al.*, 2014), 233 (Smita *et al.*, 2015), 155 (Katiyar *et al.*, 2012) et 163 (Yanhui *et al.*, 2006) chez le riz et 122 chez brachypodium (Zheng *et al.*, 2017).

Comme il n'existe pas de phylogénie publiée des gènes MYB chez le sorgho, nous avons utilisé l'étude la plus récente incluant les gènes MYB d'*Arabidopsis* comme base pour ancrer notre classification multi-espèces. Li *et al.* (2016) ont décrit 133 gènes *Arabidopsis* répartis dans 14 clades (1 à 14) ainsi que leurs liens avec les 127 FT MYB de la tomate. Cette classification des clades a été remobilisée pour prendre en compte la position des gènes *Arabidopsis* dans notre arbre phylogénique. Tous les gènes *Arabidopsis* identifiés par Li *et al.* (2016) sont inclus dans les 144 gènes MYB *Arabidopsis* identifiés par notre étude. La plupart de ces gènes conservent les mêmes liens phylogéniques entre l'arbre de Li *et al.* (2016) et le nôtre (Figure II-2B). La plupart des clades décrits par leurs travaux sont également conservées dans notre arbre. Les clades 2, 4, 5, 7, 8 et 11 ont été conservés même si les clades 5, 8 et 11 ont perdu quelques gènes *Arabidopsis* qui ont été alloués à d'autres branches de l'arbre. Les clades 1, 3, 6, 10 et 12 ont été séparés en deux sous-clades chacun, nommés A et B. Un nouveau clade, le clade 15, a été créé et contient majoritairement des gènes des anciens clades 9 et 13. Ces deux clades sont spécifiquement identifiés comme 15A et 15B. Un nouveau clade multi spécifique nommé clade 16 a été identifié et rassemble majoritairement des gènes non décrits par Li *et al.* (2016).

Au total, 23 gènes *Arabidopsis* ont des liens phylogénétiques différents entre les travaux de Li *et al.* (2016) et les nôtres. De plus, 11 nouveaux gènes *Arabidopsis* ont été inclus dans notre analyse en comparaison avec celle de Li *et al.* (2016). Neuf d'entre eux ont été alloués aux nouveaux clade 16 et les 2 autres à deux clades différents. Ces nouveaux gènes ont été

caractérisés comme étant des gènes MYB dans d'autres bases de données et 4 sont de la classe des MYB-related (Jin *et al.*, 2017). Les 135 MYB identifiés chez le sorgho sont distribués à travers les clades définis ci-dessus.

La grande taille des clades multispécifiques 4 et 15 (comptant 199 et 213 gènes respectivement) résulte de l'assemblage de différents sous-groupes définis à partir de la phylogénie d'*Arabidopsis* seule (le clade 4 correspond aux sous-groupes S13 et S16 et le clade 15 correspond à S4, S6, S7 et S15 selon Dubos *et al.* (2010)). Il est intéressant de noter que les clades 1B et 15 présentent de gros déficits de gènes monocotylédones avec seulement 33 et 26% de gènes de monocotylédones respectivement alors que les autres clades ont une moyenne de 42% de séquences de monocotylédones. Ces déficits correspondent à des diversifications extrêmes chez l'eucalyptus et le peuplier. Notre clade 1B correspond aux sous-groupes S2 et S3, ces derniers ayant été identifiés chez *Arabidopsis* comme étant impliqués dans la biosynthèse des parois. Le clade 15 rassemble les groupes *Arabidopsis* S4, S5, S6 et S7 dont certains membres sont impliqués dans la régulation de la biosynthèse des phenylpropanoïdes. A l'opposé du clade 1B, un enrichissement du clade 1A en séquences de monocotylédones est observé. Ce clade correspond au sous-groupe S1 chez *Arabidopsis* et les gènes qu'il contient ont un rôle dans la réponse aux stress abiotiques.

Parmi les 17 gènes MYB pour lesquels il existe des évidences fonctionnelles d'implication dans la régulation de la mise en place des parois secondaires, 9 appartiennent au clade 4. Huit autres FT MYB régulateurs des parois secondaires chez *Arabidopsis* appartiennent aux clades 1B (2 gènes), 12 (3 gènes) et 15 (3 gènes). Une bonne conservation du nombre des gènes entre *Arabidopsis* et le sorgho est également observée pour les MYB régulateurs des parois secondaires. Par exemple, AtMYB46 et AtMYB83 ont un seul orthologue chez le sorgho (Tableau II-2, Figure II-3B).

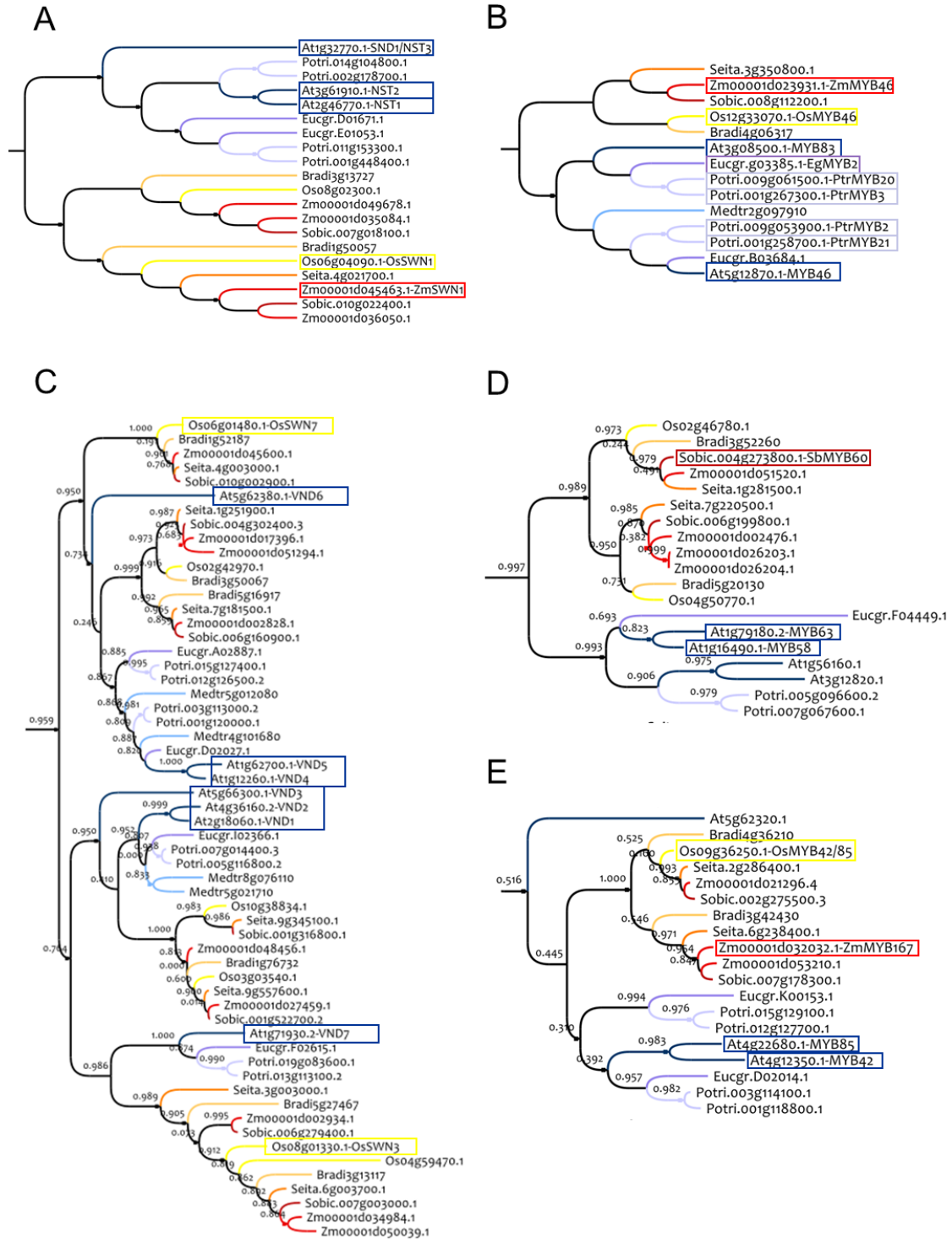


Figure II-3: Portions des arbres généraux des MYB et des NAC indiquant les homologues chez 9 espèces d'une partie des principaux régulateurs des parois secondaires. (A) homologues de AtNST1, AtNST2 et AtNST3/SND1 (Clade C) (B) homologues de AtMYB83 et AtMYB46 (Clade 4) (C) homologues de AtVND1/2/3/4/5/6/7 (Clade C) (D) homologues de AtMYB58 et AtMYB63 (Clade 1B) (E) homologues de AtMYB85 et AtMYB42 (Clade 4). Les couleurs chaudes représentent les branches monocotylédones (riz, maïs, millet des oiseaux, brachypodium et sorgho) et les couleurs froides représentent les branches dicotylédones (luzerne, eucalyptus, peuplier et Arabidopsis). Les cadres de couleurs mettent en évidence les gènes validés fonctionnellement.

3.2. Identification d'acteurs de la biosynthèse des parois par la construction de réseaux de co-expression de gènes

3.2.1. Caractérisation des clusters de gènes identifiés

Les 3 années d'expérimentation ont permis d'obtenir un peu plus de 4,7 milliards de séquences brutes avec 1,95, 1,63 et 1,39 milliards de séquences, respectivement pour 2013, 2014 et 2015. En moyenne, 5% des séquences ont été rejetées à l'étape de nettoyage par FastQC et 85% (3,96 milliards) ont été cartographiées avec succès sur le génome de référence du sorgho (v3.1). Après les filtres d'expression, un total de 20.294 gènes a été inclus dans l'analyse de co-expression. Quarante-cinq modules contenant de 32 à 2937 gènes ont été détectés mais seuls 30 modules, pour un total de 12422 gènes, présentent un seuil de corrélation de co-expression supérieur au seuil fixé et ont été analysés en détail. Parmi ces modules, 4 contiennent moins de 10 gènes, 12 entre 10 et 100 gènes, 10 entre 100 et 1000 et les 4 modules restants contiennent plus de 1000 gènes (Annexe SII-1B).

Le clustering de l'expression normalisée et moyenne de chacun des modules met en évidence des modules avec des patrons d'expression proches. Sept groupes (C1 à C7) ont été identifiés (Figure II-4). Pour illustrer les différentes propriétés des modules obtenus en termes de dynamique d'expression pendant le développement de l'entrenœud, les patrons d'expression observés pour l'entrenœud 12 en condition d'irrigation en 2013 ont été reportés afin d'avoir une première image de leur profil d'expression (Figure II-4A). Le groupe C6 présente des hauts niveaux d'expression au premier stade D1. Les groupes C1 à C3 rassemblent des gènes dont l'expression augmentent entre le stade D1 et D2 (groupe C1 et C3) et le stade D1 et D3 (groupe C2). Les groupes C5 et C7 rassemblent des gènes dont l'expression augmente en continu pendant tous les stades de prélèvement. Le groupe C4 présente une diminution de l'expression entre les stades D1 et D3 suivie d'une augmentation constante jusqu'au stade D5.

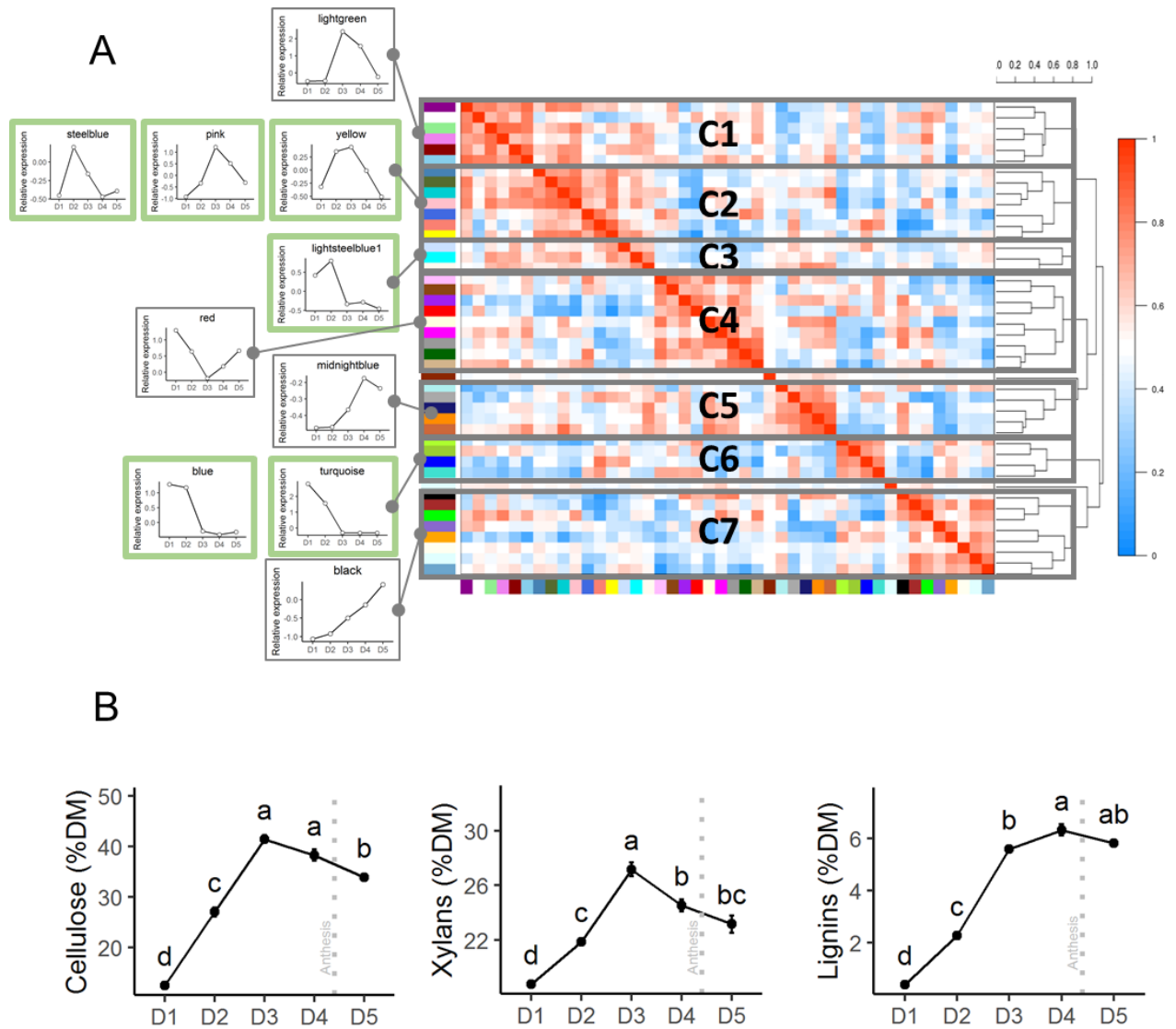


Figure II-4: Patterns d'expression des modules de co-expression identifiés par les analyses WGCNA et dynamiques d'accumulation des principaux composants pariétaux (A) Heatmap des corrélations basées sur les patrons d'expression moyens normalisés de tous les gènes de chaque module. Le clustering des modules en 7 groupes a été basé sur le dendrogramme représenté sur la droite. Les patrons d'expression obtenus pour l'expérimentation de 2013 (entre-nœud 12 en conditions irriguées) de certains modules sont présentés à gauche de la matrice de corrélation pour illustrer les patrons d'expression généraux des différents groupes de modules pendant les 5 différents stades développementaux qui ont été analysés. Les modules pour lesquels un enrichissement en gènes et GO-terms liés aux parois a été détecté sont encadrés en vert. Si les patrons d'expression présentés correspondent seulement à l'essai de 2013, les corrélations entre les modules sont calculées sur les patrons d'expression observés pendant les 3 années d'expérimentations. (B) Patrons d'accumulation des trois principaux composants pariétaux dans l'entre-nœud 12 irrigué pendant l'expérimentation de 2013. Les stades annotés avec différentes lettres présentent des valeurs moyennes significativement différentes. Les barres d'erreur correspondent à l'erreur standard basée sur 4 réplicats biologiques, la date d'anthèse est indiquée par une barre verticale grise pointillée.

Une liste d'enzymes annotées comme étant impliquées dans l'établissement des parois a été compilée à partir de résultats précédemment publiés et les 30 modules ont été testés pour un enrichissement en cette liste. Une analyse de l'enrichissement en GO-terms liés aux parois a été aussi menée pour les catégories de processus biologiques (BP) et de compartiment cellulaire (CC) issues de la Gene Ontology (<http://geneontology.org/>).

Six modules appelés « turquoise », « yellow », « pink », « blue », « steelblue » et « lightsteelblue », sont significativement enrichis en enzymes liées aux parois issus de la liste compilée pour cette étude (Tableau II-3). Les mêmes six modules sont également enrichis en GO-terms liés aux parois secondaires et aux parois en général (de 4 à 22 GO-terms de processus biologiques enrichis et de 0 à 2 GO-terms de compartiment cellulaires enrichis). Six autres modules sont enrichis avec entre 1 et 5 GO-terms de processus biologiques liés aux parois mais ne sont pas enrichis en enzymes liées aux parois (« royalblue », « green », « cyan », « darkgreen », « orangered », « darkturquoise »). Trois modules ont un enrichissement en GO-term de compartiment cellulaire paroi mais ne sont pas enrichis en enzymes liées aux parois (« saddlebrown », « lightgreen », « greenyellow »). Seuls les 6 modules enrichis à la fois en GO-terms liés aux parois et enrichis en enzymes présentes dans la liste compilée de gènes liés aux parois ont été étudiés plus en détail (Tableau II-3).

Tableau II-3: Analyse d'enrichissement des modules basée sur une liste d'enzymes impliqués dans la biosynthèse des parois et des GO-terms liés aux parois (seuls les modules avec au moins un enrichissement significatif sont présentés.)

Module	Enrichissement		Enrichissement		Nombre de gènes ³
	en aux value) ¹	enzymes parois (p-	en GO-terms ²		
			BP	CC	
Yellow	2.07*10 ⁻³⁴		22	1	1616
Blue	8.72*10 ⁻⁵		11	2	2159
Turquoise	0.00173		8	0	2855
Pink	0.0093		7	0	283
Steelblue	6.01*10 ⁻⁵		6	1	52
Lightsteelblue	1.59*10 ⁻⁵		4	2	9
Darkturquoise	0.16 ^{ns}		1	0	38
Orangered	0.33 ^{ns}		2	0	21
Darkgreen	0.49 ^{ns}		2	0	86
Cyan	0.52 ^{ns}		5	0	248
Green	0.98 ^{ns}		3	0	671
Royalblue	1 ^{ns}		3	1	63
Lightgreen	1 ^{ns}		0	1	5
SaddleBrown	1 ^{ns}		0	1	5
Greenyellow	1 ^{ns}		0	1	426

¹ ns signifie que l'enrichissement n'est pas significatif

² nombre de GO-terms liés aux parois de Processus Biologique (BP) et de Compartiment Cellulaire (CC) significativement (p-value < 5%) enrichis dans chaque module

³ nombre total de gènes dans le module considéré

Le module « yellow » présente le plus grand nombre (22) de GO-terms liés aux parois et parois secondaires avec une surreprésentation des gènes liés aux pectines, hémicelluloses, phénylpropanoïdes, à la lignine et à la cellulose (Tableau II-4). Les modules « pink » et « blue » sont aussi enrichis en GO-terms liés aux parois secondaires et plus spécifiquement en gènes liés aux lignines et aux phénylpropanoïdes. Les modules « turquoise », « steelblue » et « lightsteelblue » sont enrichis en GO-terms liés à des annotations fonctionnelles liées à la mise en place des parois cellulaires en général mais aussi aux pectines, cellulose et hémicelluloses.

En considérant l'expression moyenne des gènes qu'ils contiennent, les modules « yellow » et « pink » (groupe C2) atteignent un pic au stade D3 (dans le cadre de l'essai 2013 pour l'entre-nœud 12 en conditions irriguées). Les expressions moyennes maximales des gènes appartenant aux autres modules enrichis en GO-terms liés aux parois et aux parois secondaires interviennent plus tôt dans le développement : les modules « turquoise » et « blue » au stade D1 tandis que les modules « steelblue » et « lightsteelblue » l'atteignent au stade D2.

Les patrons d'accumulation des composants des parois de l'entre-nœud 12 en conditions irriguées au cours de l'essai en champ de 2013 sont présentés (Figure II-4B). Les compositions en hémicelluloses et celluloses atteignent un pic au stade D3, un peu plus tôt que la lignine qui est strictement spécifique des parois secondaires, au contraire des deux polysaccharides.

Tableau II-4: Enrichissement des sous-réseaux des gènes d'accroches MYB et NAC et des modules WGCNA en GO-terms liés aux parois. La colonne « Mot clé » correspond à l'agrégation de différents GO-terms décrits selon leur identifiant GO de Processus Biologique. Les cases gris foncé indiquent un enrichissement significatif (p-value < 5%). Le nombre total de gènes, le nombre total de GO-terms significativement enrichis dans le sous-réseau ou le module WGCNA considérés et l'appartenance à un module (pour les FT d'accroche) sont indiqués dans les lignes d'informations générales. La couleur des gènes d'accroche qualifie le nombre de GO-terms liés aux parois enrichis dans chaque sous-réseau : rouge pour 12 à 18 GO-terms (fortement enrichi), orange pour 6 à 12 GO-terms (moyennement enrichis) et jaune pour 3 à 6 GO-terms (légèrement enrichi). Les noms de gènes soulignés sont des MYB et des NAC jamais identifiés dans la régulation des parois secondaires.

Mot-clé	GO ID ¹	Gènes d'accroche															Modules					
		<u>SbNAC80</u>	<u>SbMYB106</u>	<u>SbMYB104</u>	SbNSTb	SbMYB83	<u>SbNAC29</u>	<u>SbMYB7</u>	<u>SbMYB92a</u>	<u>SbMYB81</u>	SbMYB60b	<u>SbNAC58</u>	<u>SbNAC8</u>	<u>SbMYB34</u>	<u>SbMYB89</u>	SbVND7a	Yellow	Pink	Blue	Sblue ²	Lsblue ³	Turquoise
SCW	GO:0009834																					
	GO:1901347																					
	GO:1901348																					
	GO:2000652																					
Lignines	GO:0046274																					
	GO:1901430																					
	GO:0009808																					
	GO:0009809																					
CW	GO:0052386																					
	GO:0016998																					
	GO:0042546																					
	GO:0042547																					
	GO:0009831																					
	GO:0009664																					

	GO:0042545																				
	GO:0009828																				
	GO:0009830																				
	GO:0071668																				
	GO:0006075																				
Xylanes	GO:0045492																				
	GO:1990937																				
	GO:0015790																				
	GO:0042732																				
	GO:0010417																				
	GO:0045493																				
	GO:0010411																				
Cellulose	GO:0030244																				
	GO:0052324																				
	GO:0010215																				
	GO:0052541																				
Pectines	GO:0010400																				
	GO:0045489																				
	GO:0010306																				
	GO:0010246																				
	GO:0010253																				
	GO:1902066																				
	GO:0045490																				
	GO:0052325																				
	GO:0048354																				
	GO:0033481																				
	GO:0019567																				

	GO:0010405																					
Phenylpr- opanoïdes	GO:0033587																					
	GO:0019632																					
	GO:0009699																					
	GO:0009805																					
	GO:0009094																					
	GO:0006559																					
	GO:0080088																					
	GO:0009812																					
Flavonoides	GO:0031539																					
	GO:1900384																					
	GO:0010023																					
	GO:0009715																					
	Développeme nt des vaisseaux	GO:0090058																				
GO:0010089																						
GO:0010222																						
Cires	GO:0010345																					
	GO:0010143																					
Informations générales																						
Nombre of gènes considérés		66	63	71	51	70	85	59	83	82	39	30	25	67	28	41	161	28	215	52	9	285
Nombre de GO-terms significatifs		31	27	31	32	31	16	13	19	15	35	28	16	14	14	15	92	75	105	28	8	205
Module		yellow	yellow	yellow	yellow	yellow	blue	blue	yellow	blue	pink	steelblue	yellow	blue	yellow	yellow						

¹ Les GO-terms associés aux GO ID sont disponibles en Annexe SII-2

² Module Steeblue

³ Module Lighsteeblue

3.2.2. Identification des « hub genes »

Au sein de chacun des modules enrichis en gènes et en ontologie de « parois » et « parois secondaires », les gènes très connectés ont été identifiés en utilisant le paramètre de connectivité intra-module. Les 10 meilleurs gènes ont été reportés sur le Tableau II-5. A l'exception du module « lightsteeblue » qui contient un faible nombre de gènes, les 10 meilleurs gènes sont les mêmes que nous considérons l'ensemble du module sans seuil de corrélation ou avec, ce qui révèle que les gènes ayant le plus de connections dans le module sont aussi ceux ayant les plus fortes corrélations avec les autres gènes du module (Tableau II-5).

Certains de ces gènes sont annotés comme étant des FT selon leur annotation GO (GO:0003700). Il s'agit de, Sobic.002G260800, une protéine de la famille des ERF dans le module steeblue, et Sobic.004G237300, un FT TCP du module blue et deux protéines appartenant à la famille des ERF, Sobic.003G148600 et Sobic.010G080400, dans le module « lighsteeblue ».

En plus de ces FT, un gène NAC et un gène MYB sont présents dans les 10 meilleurs « hub genes » du module « pink » même s'ils ne sont pas annotés comme tels au niveau des GO-terms. Il s'agit de Sobic.003G035100 (SbNAC75a) qui est un homologue proche de NAC075, un régulateur putatif de VND7 lui-même étant un régulateur principal des parois secondaires chez *Arabidopsis* (Fujiwara et Mitsuda, 2016; Yamaguchi *et al.*, 2011) et de Sobic.001G110900 qui est un homologue de AtMYB52/54, un autre régulateur des parois secondaires (Zhong *et al.*, 2008).

Le FT TCP du module « blue » est annoté comme impliqué dans le métabolisme des flavonoïdes. D'autres gènes parmi les 10 meilleurs « hub-Genes » de ces modules ont des GO-terms liés aux parois et aux parois secondaires (8 gènes au total) et ont tous, sauf le module « steeblue », au moins un gène de leur « top 10 » annotés avec une ontologie liée aux parois ou aux parois secondaires.

En plus des 10 meilleurs gènes, les 10% des meilleurs gènes des six modules « parois » ont été récupérés pour des analyses complémentaires. Parmi ces 10% de meilleurs gènes, le module « yellow » a le plus de gènes ayant des GO-terms parois ou parois secondaires (31) mais ne contient pas de FT selon les GO-terms. Cependant, nous notons que SbVND7a fait partie des 10% des meilleurs Hub-Genes du module « yellow ». SbVND7a est donc un gène crucial du réseau des gènes du module « yellow ». Les listes des meilleurs 10% de gènes des

modules « blue », « turquoise » et « pink », les plus connectés aux autres sont composées d'entre 5 et 10% de gènes ayant des GO-terms liés aux parois et parois secondaires. Elles ne comprennent ni FT, ni NAC, ni MYB, pour le module « blue » mais 3 MYB pour le module « turquoise », aucun n'étant un homologue proche des gènes validés dans d'autres espèces pour leur rôle dans la régulation des parois secondaires. Il y a également deux FT GATA et WRKY dans les 10% des meilleurs gènes du module « pink ».

Tableau II-5: Meilleurs « hub genes » dans les six modules liés aux parois. Les « hub genes » ont été identifiés en utilisant la connectivité intramodule. Pour chacun des modules, les meilleurs 10 « hub gene », leurs rôles putatifs et leur annotation avec un ou plusieurs GO-terms liés aux parois sont indiqués. Les nombre des 10% meilleurs « hub genes » de chaque module et le nombre de ces gènes associés à un ou plusieurs GO-terms liés aux parois sont également indiqués

Module	Gene ID	Role putatif (Phytozome v12 « best hit » Arabidopsis et riz)	GO-terms associés aux parois ¹	Nombre de gènes dans le top 10%	Nombre de gènes avec GO-terms parois dans le top 10%
blue	Sobic.001G502900	HR-like lesion-inducing protein-related		216	13
	Sobic.003G011100	VIRB2-interacting protein 1 / reticulon-like			
	Sobic.003G223100	Polygaracturonase (pectin lyase-like)			
	Sobic.004G237300	TCP family transcription factor	x		
	Sobic.004G323100	Cysteine rich secretory protein / defense			
	Sobic.004G333500	Proteasome maturation factor UMP1			
	Sobic.006G046200				
	Sobic.007G180200				
turquoise	Sobic.008G035400	Stress responsive A/B Barrel Domain		283	15
	Sobic.008G051900	ATPase subunit			
	Sobic.001G365266	Zing-finger type protein			
	Sobic.002G279400	Protein of unknown function			
	Sobic.004G063600	Vacuolar import/degradation, Vid27-related protein			
	Sobic.004G269100	PRONE (Plant-specific Rop nucleotide exchanger)	x		
	Sobic.006G124400	ATP binding protein / DNA-directed DNA polymerase			
	Sobic.006G241800	Leucine-rich repeat protein kinase family protein			
	Sobic.007G201900	Protein of unknown function			
	Sobic.009G172950	Minichromosome maintenance family protein / DNA helicase			
	Sobic.010G077000	Leucine-rich receptor-like protein kinase family protein	x		

pink	Sobic.010G129700	microtubule-associated protein 65-2			
	Sobic.001G110900	Homolog of AtMYB52/54	x	29	3
	Sobic.001G131400	Thioesterase superfamily protein			
	Sobic.001G361300	CDPK-related kinase 1			
	Sobic.001G372000	F-box et tubby domain containing protein			
	Sobic.001G482100	Plant protein of unknown function			
	Sobic.002G200700	Protein of unknown function (DUF1666)			
yellow	Sobic.003G035100	Homolog of AtNAC075	x		
	Sobic.003G317500	lysophosphatidyl acyltransferase 5			
	Sobic.004G102400	RING/FYVE/PHD-type zinc finger family protein			
	Sobic.006G145901	beta glucosidase 46 / monolignol beta-glucoside homolog			
	Sobic.001G038300	TRICHOME BIREFRINGENCE-LIKE 33		161	31
	Sobic.001G063500	FASCICLIN-like arabinogalactan protein 17 precursor			
	Sobic.002G128800	Protein of unknown function			
	Sobic.002G252000	Chitinase family protein	x		
	Sobic.003G266400	ENTH/ANTH/VHS Traffichink pathway / clathrin assembly protein			
	Sobic.004G111100	GDSL-like Lipase/Acylhydrolase superfamily protein			
steelblue ²	Sobic.004G221300	Glucose-6-phosphate/phosphate et phosphoenolpyruvate antiporter	x		
	Sobic.005G194900	Phosphoserine phosphatase			
	Sobic.008G003600	Plant protein of unknown function			
	Sobic.010G022300	Transmembrane receptor			
	Sobic.001G359700	DUF679 domain membrane protein 2		<10	0
	Sobic.002G260800	Integrase-type DNA-binding superfamily protein AP2 domain			
	Sobic.004G310000	Lipase / alpha/beta-Hydrolases superfamily protein			
	Sobic.006G232500	Nuclease			
	Sobic.006G244400	Nuclease			
	Sobic.007G172100	Cysteine protease			
	Sobic.009G257400	alpha-vacuolar processing enzyme			

	Sobic.010G078200			
	Sobic.010G133866			
	Sobic.010G158200	Protein of unknown function		
lightsteelblue ³	<i>Sobic.001G080100</i>		<10	2
	<i>Sobic.001G304201</i>	<i>LTPL141 - Protease inhibitor/seed storage/LTP family protein precursor</i>		
	<i>Sobic.003G148600</i>	<i>Integrase-type DNA-binding superfamily protein AP2 domain</i>		
	<i>Sobic.003G432700</i>	<i>Curculin-like (mannose-binding) lectin family protein</i>		
	Sobic.007G146200	Plant invertase/pectin methylesterase inhibitor superfamily	x	
	<i>Sobic.007G214600</i>	<i>Calcium-binding EF-hand family protein</i>		
	<i>Sobic.009G152600</i>	<i>pyrophosphorylase 3</i>		
	Sobic.009G173700	Expansin	x	
	<i>Sobic.010G080400</i>	<i>Integrase-type DNA-binding superfamily protein AP2 domain</i>		
	<i>Sobic.007G214400</i>	<i>Calcium-binding EF-hand family protein</i>		

¹ “x” signifie que le gène a au moins une annotation GO-term liée aux parois

² comme il n’y a que 52 gènes dans le module « steeblue », les informations des meilleurs 10% sont équivalentes à celles des 10 meilleurs gènes

³ comme il n’y a que 9 gènes au-dessus du seuil de corrélation dans le module « lightsteelblue », les gènes reportés dans les meilleurs 10 « hub genes » ne considèrent pas le seuil de corrélation. Les deux gènes qui ne sont pas en italique sont les deux gènes au-dessus du seuil de corrélation

3.2.3. Résultats de l'approche « gènes d'accroches »

Parmi les 257 gènes SbMYB et SbNAC inclus dans l'analyse phylogénique, 83 MYB et 86 NAC n'ont pas été considérés dans les analyses de réseau de gènes à cause de leur faible expression ou de l'absence de forte corrélation avec d'autres gènes. Cela concerne certains homologues sorgho de gènes clés de la régulation des parois secondaires qui ont été exclus à cause de leur niveau d'expression (SbVND4-6c, SbVND7b, SbVND1-3a,b, SbNAC75c,d,e, SbVNI2a,d,c et SbMYB20/43c, SbMYB69a,b) ou de leurs niveaux de corrélations en dessous du seuil choisi (SbMYB20/43a, SbMYB42b, SbMYB4b, SbNAC75b, SbVNI2f et SbMYB60).

La plupart des 52 MYB et 36 NAC inclus dans les réseaux de co-expression font partie du module « brown » (11 MYB et 14 NAC), du « blue » (10 MYB et 3 NAC) et du « yellow » (13 MYB et 9 NAC) (Figure II-5). Le reste des gènes MYB et NAC se trouvent dans les modules « turquoise », « steelblue » et « pink » mais aussi dans les modules « red », « green » et « cyan. » Aucun gène MYB et NAC n'appartient au module « lighsteelblue » qui ne contient que 9 gènes au seuil de corrélation que nous considérons.

En dehors d'un gène dans le module « green », l'ensemble des 30 gènes MYB et 14 gènes NAC dont les sous-réseaux sont enrichis en GO-terms liés aux parois appartiennent aux modules qui présentent des enrichissements en GO-terms et gènes de structure liés aux parois primaires et secondaires.

Parmi ces sous-réseaux associés aux FT enrichis en GO-terms liés aux parois, 22 se trouvent dans le module « yellow », 10 dans le module « blue », 7 dans le module « pink », 3 dans le « turquoise » et 1 dans le « steelblue » (Figure II-5). Le détail

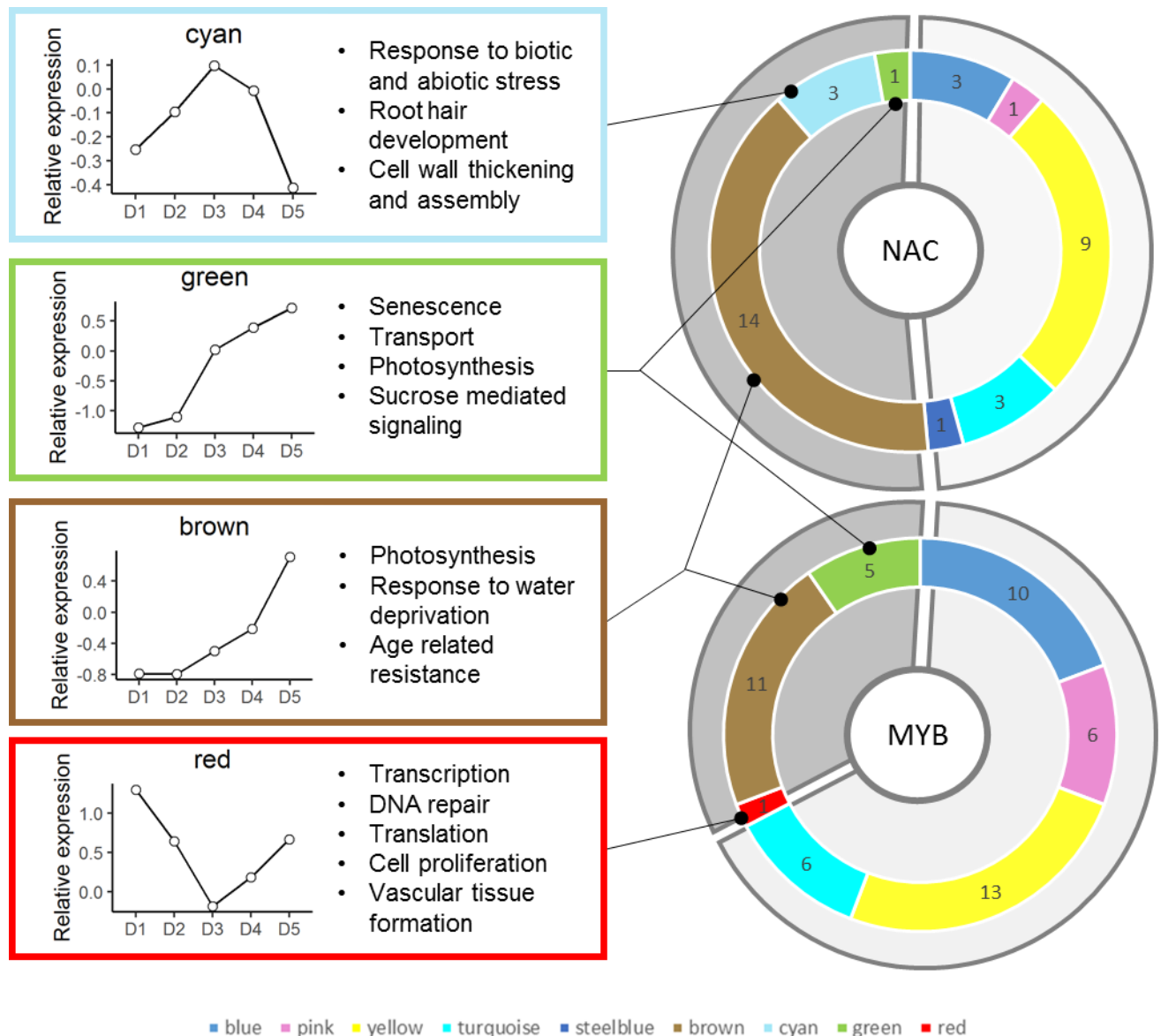


Figure II-5: Distribution et nombres de NAC et MYB chez le sorgho dans les différents modules. Les modules enrichis en GO-terms liés aux parois sont mis en évidence en gris clair. Les modules qui ne sont pas directement reliés aux parois (i.e. « brown », « cyan », « green » et « red ») sont mis en évidence en gris foncé. Les principaux enrichissements en GO-terms des modules en gris foncé sont indiqués dans les cadres à gauche ainsi que le patron d'expression moyen de leurs gènes dans l'expérimentation 2013.

Nous avons utilisé 88 gènes MYB et NAC comme des gènes d'accroche pour identifier les gènes avec lesquelles ils sont fortement co-exprimés. Nous avons choisi d'en extraire la liste des 15 gènes les plus co-exprimés par gène d'accroche et d'itérer cette étape en utilisant ces 15 gènes comme accroches à leur tour. Cela permet d'obtenir, pour chaque MYB et NAC considéré, un réseau à deux niveaux théoriquement composé d'un maximum de 226 gènes ($1+15 \times 15$). Les « sous-réseaux » générés contiennent entre 22 et 102 gènes avec un nombre moyen de 65,8 gènes. L'enrichissement en GO-terms de ces sous-réseaux a été caractérisé en se focalisant principalement sur les GO-terms liés aux parois.

Nous avons caractérisé plus en détails les sous-réseaux qui sont enrichis en 3 à 18 GO-terms liés aux parois, ce qui représente les sous-réseaux de 39 FT, et les avons classifiés en trois catégories d'enrichissement : « fortement enrichi » entre 12 et 18 GO-terms, « moyennement enrichi » entre 6 et 12 et « légèrement enrichi » entre 3 et 6 (dont une partie est représentée dans le Tableau II-4). Les 19 sous-réseaux fortement enrichis appartiennent au module « yellow ». Deux moyennement enrichis appartiennent au module « yellow » également et 5 moyennement enrichis appartiennent au module « Blue » et 4 au module « pink ». Trois sous-réseaux légèrement enrichis appartiennent au module « yellow », 3 autres au module « blue », et enfin les modules « pink », « steelblue » et « turquoise » en comptent chacun un.

L'enrichissement des sous-réseaux liés à chacun des orthologues SbMYB et SbNAC de MYB et NAC validés pour leur rôle dans la régulation de la mise en place des parois secondaires chez d'autres espèces ont été analysés. Dans la plupart des cas les sous-réseaux associés aux gènes homologues du sorgho sont enrichis en GO-terms liés aux parois. Pour des contraintes d'espace, seule une partie des enrichissements des sous-réseaux des FT d'accroche est représentée dans le Tableau II-4.

En commençant par les gènes au sommet de la voie de régulation ; les sous-réseaux de SbNSTb, SbNSTa, SbSNDb et SbSNDa sont enrichis en GO-terms spécifiques de la synthèse des lignines et en GO-terms liés à la mise en place générale des parois (l'un d'eux est représenté à titre d'exemple dans le Tableau II-4).

Les homologues « sorgho » des gènes Arabidopsis VND1 à VND6, qui sont impliqués dans le dépôt spécifique des parois secondaires dans les cellules vasculaires, n'ont pas passé les filtres de sélection d'expression et de corrélation ou leur sous-réseau n'est que faiblement ou pas du tout (SbVN4a dans le module « turquoise » et SbVND4-6b dans le module « blue ») enrichis en GO-terms liés aux parois. Concernant le gène d'Arabidopsis VND7, le seul parmi les deux

orthologues sorgho à passer nos filtres de sélection est SbVND7a. Son sous-réseau est légèrement enrichi en GO-terms liés aux cires présentes dans les parois (subérine, cutine), en composés phénoliques et en pectine (Tableau II-7).

Le sous-réseau d'un des 7 orthologues de VNI2 régulateur des parois secondaires chez *Arabidopsis* et interagissant avec VND7, SbVNI2g, est un des deux sous-réseaux NAC avec celui de SbSNDa les plus enrichis en GO-terms liés aux parois.

L'autre niveau majeur de régulation des parois secondaires chez *Arabidopsis*, le maïs et le riz correspond au MYB46 qui est redondant fonctionnellement avec son paralogue MYB83 chez *Arabidopsis*. Chez le sorgho, il existe un seul orthologue direct de ces deux gènes, SbMYB83, (Tableau II-1) qui appartient au module « yellow ». Son sous-réseau est fortement enrichi en 12 GO-terms liés à la mise en place générales des parois (xylane, cellulose) mais aussi spécifiques des parois secondaires.

SbMYB60 et son paralogue SbMYB60b sont des orthologues de AtMYB58 et AtMYB63. SbMYB60 a été validé fonctionnellement pour son implication dans la voie de biosynthèse des lignines chez le sorgho (Scully *et al.*, 2016b, 2017) mais ne passe pas notre seuil de corrélation. SbMYB60b est, lui, considéré dans notre analyse finale et appartient au module « pink ». Son sous-réseau est moyennement enrichi en une ontologie liée à la lignine, 3 GO-terms liés aux phénylpropanoïdes et 2 GO-terms liés à la biosynthèse des flavonoïdes.

Les gènes SbMYB homologues des régulateurs des parois secondaires en aval (Figure I-20) ont leurs sous-réseaux, tout comme les homologues des régulateurs de premier niveau, fortement enrichis en GO-terms liés aux parois. Les sous-réseaux de SbMYB4a, un orthologue du groupe de MYB répresseurs des parois secondaires comprenant AtMYB4, AtMYB7 et AtMYB32, et SbMYB61b, un orthologue de AtMYB61, sont tous deux fortement enrichis en GO-terms liés aux parois, respectivement avec 18 et 17 GO-terms.

Nous avons aussi détecté des enrichissements dans les sous-réseaux de gènes SbMYB et SbNAC dont les homologues chez d'autres espèces ne sont pas identifiés comme impliqués dans la mise en place des parois secondaires (Tableau II-4, Tableau II-6). Parmi eux, les sous-réseaux de 4 SbMYB et un SbNAC sont fortement enrichis en GO-terms liés aux parois, 8 SbMYB et un SbNAC sont moyennement enrichis et 3 SbMYB et 3 SbNAC sont légèrement enrichis.

Les sous-réseaux de SbNAC80, SbMYB17, SbMYB40, SbMYB104, et SbMYB106 sont fortement enrichis en GO-terms liés aux parois. Ceux de SbNAC29, SbMYB7, SbMYB43, SbMYB53, SbMYB76, SbMYB81, SbMYB92, SbMYB92a, b, et c sont moyennement enrichis et ceux de SbNAC16, SbNAC28, SbNAC58, SbMYB34, SbMYB55 et SbMYB89 le sont légèrement (Tableau II-4).

Tous les sous-réseaux de FT fortement enrichis appartenant au module « yellow » sont enrichis en GO-terms “plant-type SCW biogenesis” (biogénèse des parois secondaire végétales) ainsi qu’en au moins un GO-term lié à la lignine (“metabolic”, “catabolic”, “biosynthesis”, métabolisme, catabolisme ou/et biosynthèse de la lignine), “xylan biosynthetic process” (biosynthèse des xylanes) et “cellulose biosynthetic process” (biosynthèse de la cellulose). Aucun n’est, par contre, enrichi en GO-terms liés aux flavonoïdes et aux phénylpropanoïdes.

Les 11 sous-réseaux moyennement enrichis en GO-terms liés aux parois font partie des modules « blue », « pink » et « yellow » et seuls 2 d’entre eux sont enrichis en GO-terms spécifiques des parois secondaires. Il est intéressant de noter que les 4 qui appartiennent au module « pink » sont enrichis en GO-terms liés aux flavonoïdes et aux phénylpropanoïdes.

Aucun des sous-réseaux de FT légèrement enrichis en GO-terms liés aux parois et qui font partie des modules « blue », « steelblue » et « yellow », ne sont spécifiquement enrichis en GO-terms liés aux parois secondaires. Parmi eux, les sous-réseaux de SbNAC8, SbMYB89 et SbVN7a sont enrichis en GO-terms liés aux cires.

Tableau II-6: Gènes MYB et NAC du sorgho sans homologie avec des FT régulateurs des parois validés fonctionnellement chez d'autres espèces et identifiés pour leur rôle potentiel dans la mise en place des parois.

Famille	Clade	Gene ID	Nom	Module	Gènes Arabidopsis les plus proches	FT Arabidopsis dans le même clade ¹	Branche spécifique des monocotylédones ²	Liens phylogénétiques ³
NAC	G	Sobic.006G118200	SbNAC80	yellow	At1g03490 At1g28470 At1g60240 At1g60280 At1g60300 At1g60340 At1g60350 At1g60380 At3g04430	SND2, SND3, NAC075	x	
NAC	E	Sobic.002G290800 Sobic.001G529200	SbNAC29 SbNAC18	blue blue	At3g17730 At5g17260 At1g65910 At3g03200			
NAC	B	Sobic.004G219500	SbNAC58	steelblue	At4g28530			
NAC	A	Sobic.001G163800	SbNAC8	yellow	At5g61430			
MYB	4	Sobic.003G183800 Sobic.009G216900 Sobic.003G270300	SbMYB40 SbMYB106 SbMYB43	yellow	At1g63910	Neuf MYB régulateurs des parois	x	paralogues
MYB	11	Sobic.009G103900; Sobic.002G141300; Sobic.004G070900	SbMYB104; SbMYB17; SbMYB53	yellow; yellow; blue	At4g37260 At2g23290 At5g67300 At3g50060			paralogues
	15A	Sobic.009G148100;	SbMYB92a,	yellow;	At1g22640	MYB4, MYB7,	x	paralogues

MYB		Sobic.005g062000; Sobic.008G055700; Sobic.008G055800	SbMYB92b, SbMYB92c, SbMYB92	pink; pink; pink	At1g35515 At2g16720 At4g09460 At4g34990 At4g38620	MYB32	
MYB	12B	Sobic.003G129000	SbMYB34	blue	At1g17950 At1g26780 At1g69560 At1g73410 At3g29020 At4g33450 At5g17800 At5g39700	MYB69, MYB52, MYB54	x
MYB	3B	Sobic.007G136500 Sobic.004G191800	SbMYB81 SbMYB55	blue	At3g01140 At5g15310		paralogues
MYB	2	Sobic.007G224800	SbMYB89	yellow	At1g34670 At5g65230 At5g10280		
MYB	1B	Sobic.001G219200	SbMYB7	blue	At1g06180 At2g31180 At3g23250	MYB58, MYB63	

¹ Nom de gènes d'Arabidopsis avec des évidences fonctionnelles de rôle dans la regulation de la mise en place des parois secondaires

² x signifie que le gène sorgho considéré appartient à une branche phylogénétique sans proche homologie avec les dicotylédones ou avec Arabidopsis

³ Liens d'homologies entre les gènes sorghos considérés s'ils sont proches

La compilation des FT dont les sous-réseaux sont les plus enrichis en GO-terms liés aux parois (les sous-réseaux fortement enrichis pour le module « yellow » et les sous-réseaux moyennement enrichis pour le module « pink ») ainsi que les sous-réseaux des FT dont les homologues sont déjà connus chez d'autres espèces montrent de nombreuses interconnexions de co-expression (Figure II-7A et B). Trente-trois enzymes issues de la liste des gènes impliqués dans les parois sont présentes parmi les 207 gènes connectés dans la Figure II-7A qui représente une partie du module « yellow » centré sur les FT les plus intéressants (tels que désignés ci-dessus ; leurs sous-réseaux sont les plus enrichis en GO-terms liés aux parois).

Deux CAFFEYOYL-COA 3-O-METHYLTRANSFERASE (CCoAMT) et une PHENYLALANINE AMMONIA-LYASE (PAL), qui sont des gènes de structure clés de la voie de biosynthèses des lignines, sont, eux, connectées dans le réseau de 59 gènes représenté dans la Figure II-7B (portion du module « pink » focalisée sur ses FT les plus intéressants). Il est intéressant de noter que 138 et 44 gènes dans ces portions de module « yellow » et « pink » n'ont jamais été reliés à la synthèse des parois jusqu'à présent dans la littérature (Figure II-7A et B).

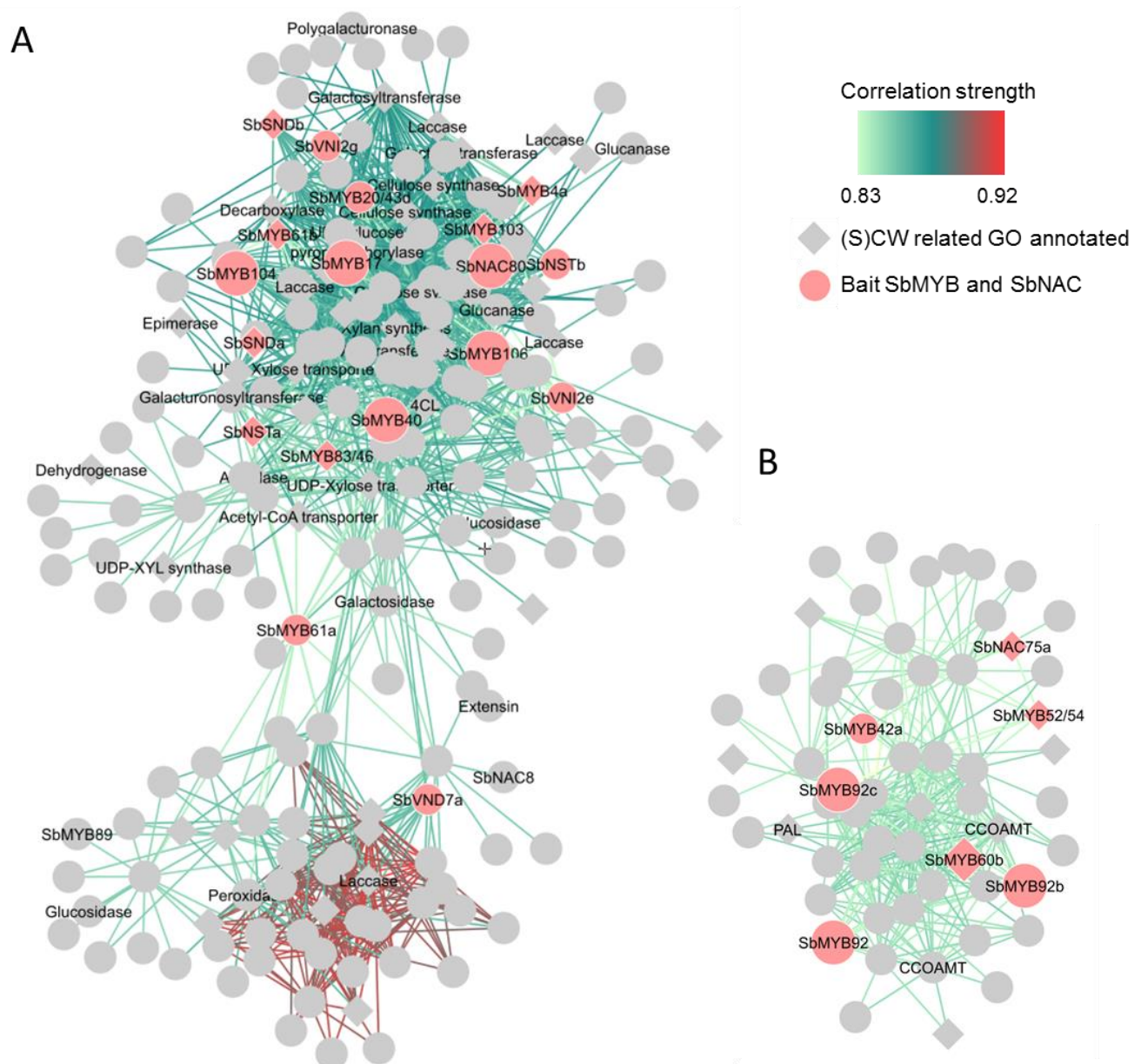


Figure II-7: Réseaux compilés des sous-réseaux des homologues chez le sorgho des NAC et des MYB validés fonctionnellement chez *Arabidopsis* pour leur rôle dans la régulation des parois secondaires et des nouveaux SbMYB et SbNAC sans annotation fonctionnelle identifiés par l'analyse de co-expression de gènes. Les nouveaux FT reportés sont ceux dont le sous-réseau présente le plus d'enrichissement en GO-terms liés aux parois. (A) Dans le module jaune, les nouveaux FT sont « fortement enrichis » en GO-terms liés aux parois. (B) Dans le module rose, les nouveaux FT sont « moyennement enrichis » en GO-terms liés aux parois. En orange clair sont indiqués les SbMYB et SbNAC utilisés en gènes d'accroche pour extraire les sous-réseaux compilés dans cette figure. Les cercles plus grands correspondent aux SbMYB et SbNAC identifiés comme associés aux parois dans cette étude. Les carrés correspondent aux gènes ayant au moins un GO-term lié aux parois. La corrélation est présentée selon un gradient de couleur de vert clair (plus faibles corrélations) à rouge (plus fortes corrélations). Les noms des enzymes sont issus de la liste des enzymes impliqués dans les parois construite à partir de la littérature.

3.3. Gènes mis en évidence par des analyses de génétique d'association

3.3.1. 58 « hotspots » liés à la qualité de la tige ont été identifiés

La détection des associations au sein de chaque essai et sur les valeurs de BLUP issues de plusieurs essais a permis d'identifier 83 locus SNP uniques associés à des caractères de qualité de la biomasse (méta-trait : paroi, sucres solubles et digestibilité). Certains ont été identifiés plusieurs fois et dans l'ensemble : 31 ont été détectés sur la base des valeurs de blup (qui prend en compte différents essais), 27 avec les analyses menées sur des essais uniques et 25 qui ont été détectés à la fois avec les valeurs de blup pour le blup et sur les essais uniques. Les effets de ces SNP ont été associés à des caractères liés au contenu des parois (94), à la digestibilité (17) et au contenu en sucres (3). Ils sont localisés dans 58 « hotspot » d'intérêt sur le génome dont 18 sont constitués d'au moins deux loci SNP. Onze hotspots sont associés à différents méta-trait, 9 d'entre eux étant des associations des traits de contenu en paroi et digestibilité et 2 associant les traits de contenus en sucres et en parois. Les hotspots associés à un seul meta-trait sont 41 à être associés au contenu en parois, 5 à la digestibilité et 1 au contenu en sucres.

Deux cent cinquante-sept gènes sont situés à proximité (sur une fenêtre de +/- 20kb autour des extrémités de la zone) de ces hotspots impactant la variabilité de caractères liés à la composition de la biomasse et des parois ce qui représente moins de 5 gènes par zone en moyenne. Plusieurs zones ne contiennent aucun gène, c'est le cas d'une zone associée aux caractères de paroi et de 3 zones associées à la fois à des caractères de paroi et de digestibilité. Selon les GO-terms associés à ces gènes, ils seraient impliqués dans 53 processus biologiques et 72 fonctions moléculaires différents. Ce sont donc des gènes aux rôles variés allant du métabolisme des lipides à l'expression et la synthèse des protéines.

Six gènes sont présents dans des listes de gènes issues de la littérature associées aux parois chez le sorgho. Il s'agit d'enzymes impliquées dans la modification et la biosynthèse des polysaccharides de structure. Deux xyloglucan galactosyltransférases, une xyloglucan fucosyltransférase, une galactosyltransférase, une xyloglucan endotransglucosylase et une pectin acétyl estérase. D'autres gènes assurent des fonctions qui pourraient être reliées à la qualité de la biomasse comme une pectine acyl estérase, une serine acétyltransférase, trois alpha-amylases et deux enzymes appartenant aux familles des chalcone et stilbène synthases. Plusieurs FT des familles GATA, TCP, Zinc finger, E2F, bhlh et MADS-box sont également

présents. Deux FT de la famille des NAC et MYB sont présents. Il s'agit des homologues chez le sorgho de AtMYB105 et AtVND4.

Il serait bien d'avoir un tableau de synthèse indiquant pour les différents zones quels sont les gènes sous jacents qui ont été identifiés (ceux qui font partie des listes de candidats de la biblio et des NAC / MYB)

3.3.2. Comparaison des listes de gènes candidats issues des analyses phylogénétiques, des réseaux de gènes mis en évidence et des régions chromosomiques détectées contrôlant la variabilité de la qualité de la biomasse

Parmi les 257 gènes situés au sein des zones chromosomiques associées aux caractères de la tige (digestibilité, sucre, paroi végétale), 112 sont également présents dans les modules associés aux parois et aux parois secondaires mis en évidence par l'analyse de réseau de co-expression. En particulier 24 gènes dans le module « turquoise », 23 gènes dans le « blue », 22 dans le « yellow » et 1 dans le « pink ». Les deux autres modules identifiés comme étant importants dans la mise en place des parois, « lightsteelblue » et « steelblue », ne contiennent donc pas de gènes situés dans les zones associées au phénotype de la tige et 54 gènes se trouvent dans d'autres modules non reliés à la mise en place des parois par les analyses d'enrichissement des réseaux de co-expression de gènes.

Parmi ces gènes, deux gènes dans chacun des modules (sauf le module « pink » qui n'en contient aucun) font partie des 10% de gènes les plus connectés (« hub genes ») mais aucun n'est dans les 10 meilleurs des gènes les plus centraux (Tableau II-7). Pour le module « blue » il s'agit d'un gène non annoté et d'une lactate/malate deshydrogénase, pour le module « turquoise », il s'agit d'un gène non annoté et d'une copper oxidase et pour le module « yellow » il s'agit de deux xyloglucane galactosyltransférases (GalT).

Tableau II-7 : Quatorze gènes retenus présents dans les modules de co-expression liés aux parois et présents les zones d'intérêt GWAS pour les méta-trait d'intérêts CW (contenu en parois), AD (digestibilité) et SJ (contenu en sucres solubles) ou étant un MYB ou un NAC. En plus de ces critères, ces gènes appartiennent à la liste de « hub gene » de leur module (leur centralité dans leur module est indiquée par le rang) ou à un sous-réseau de régulation d'un MYB ou d'un NAC ou à la liste d'enzymes impliquées dans la biosynthèse des parois. Les homologues Arabidopsis (At best hit) et riz (Os best hit) ainsi que la fonction de ces gènes sont indiqués

Nom du gène	Méta-trait d'association	Module de co-expression	Rang "hub gene"	Intégration dans un sous-réseau MYB ou NAC	Enzyme de la biosynthèse des parois	At best hit	Os best hit	Fonction putative
Sobic.009G240700	AD	blue	140	non	non	AT1G53240	LOC_Os01g46070	mMDH1;Lactate/malate dehydrogenase family protein
Sobic.006G203700	CW	blue	162	non	non	AT1G64850	LOC_Os04g51250	Calcium-binding EF hand family protein
Sobic.002G194500	AD+CW	blue	NA	SbMYB7	oui	AT1G14720	LOC_Os09g23220	XTH28,XTR2;xyloglucan endotransglucosylase/hydrolase 28
Sobic.010G003100	CW	turquoise	200	non	non	AT4G12420	LOC_Os06g01490	SKU5;Cupredoxin superfamily protein;monocopper oxidase
Sobic.009G208900	CW	turquoise	215	Sobic.003G14510	non	AT4G12540	LOC_Os05g45580	
Sobic.009G208900	CW	turquoise	215	non	non	AT4G12540	LOC_Os05g45580	
Sobic.008G171900	CW	turquoise	NA	non	non	AT1G72250	LOC_Os12g42160	Di-glucose binding protein with Kinesin motor domain
Sobic.010G002900	CW	turquoise	NA	non	non	AT1G12260	LOC_Os06g01480	NAC007;VND4
Sobic.004G308600	CW+SJ	turquoise	NA	non	oui	AT2G03220	LOC_Os02g52640	ATFUT1,FT1,MUR2;fucosyltransferase 1;Xyloglucan fucosyltransferases (MUR2)
Sobic.004G051300	CW	yellow	NA	non	oui	AT1G05170	LOC_Os02g06840	Galactosyltransferase family protein
Sobic.009G220100	CW	yellow	13	SbMYB17	oui	AT5G61840	LOC_Os01g70200	GUT1;Exostosin family protein; Xyloglucan galactosyltransferases (MUR3)
Sobic.009G220200	CW	yellow	24	SbMYB17	oui	AT1G27440	LOC_Os01g70190	GUT2,IRX10;Exostosin family protein;Xyloglucan galactosyltransferases (MUR3)
Sobic.008G171600	CW	pink	NA	SbMYB52/54	non	AT3G59030	LOC_Os12g42130	ATTT12;MATE efflux family protein
Sobic.007G137101	AD+CW	non inclus	NA	NA	non	AT1G69560	LOC_Os08g33800	LOF2;MYB105

Ces deux derniers gènes font également partie du sous-réseau de co-expression le plus corrélé avec SbMYB17 qui est un nouveau candidat fort pour être impliqué dans la régulation de la mise en place des parois secondaires chez le sorgho. En effet, le sous-réseau de SbMYB17 est très enrichi en gènes impliqués dans l'établissement des parois. En plus de ces deux gènes du module « yellow », deux gènes du module « turquoise », un gène du module « pink » et un gène du module « blue » sont inclus dans les sous-réseaux de FT MYB (mais aucun NAC) (Tableau II-7). Sobic.008G171600 (un transporteur MATE efflux appartenant au module « pink ») fait partie des gènes dont l'expression est très corrélée à SbMYB52/54 un orthologue de gène majeur de la régulation des parois secondaires chez Arabidopsis. Sobic.002G194500 (une xyloglucan endotransglucosylases/hydrolases appartenant au module « blue ») fait partie du sous-réseau de SbMYB7, également un bon candidat de la régulation des parois chez le sorgho. SbMYB7 est proche, sans en être un orthologue direct, des gènes AtMYB58 et AtMYB63, régulateurs du dépôt des lignines chez Arabidopsis. Sobic.008G171900 et Sobic.009G208900 (une protéine liée aux microtubules et une protéine sans rôle connu appartenant au module « turquoise ») font partie du sous-réseau Sobic.003G14510, un MYB qui n'est pas un gène mis en évidence comme étant impliqués dans la mise en place des parois, ni par notre analyse de réseau de co-expression de gènes, ni par ses liens phylogénétiques avec des acteurs connus chez d'autres espèces (Tableau II-7).

Concernant les FT MYB et NAC, MYB105 (Sobic.007G137101) n'est pas considéré dans les analyses de co-expression à cause de sa faible expression. Par contre l'homologue de VND4 (Sobic.010G002900) appartient au module « turquoise » qui est relié à la mise en place des parois. Cependant, malgré son appartenance à la famille VND, il n'est pas considéré comme un candidat fort de la régulation des parois secondaires chez le sorgho.

4. Discussion

4.1.Mise à jour des classifications phylogénétique des FT MYB et NAC dans l'objectif de faciliter l'identification des régulateurs principaux de la mise en place des parois secondaires

Les FT des familles MYB et NAC remplissent des fonctions variées à travers leur implication dans les réseaux de régulation qui contrôlent de développement, le métabolisme et la réponse aux stress biotiques et abiotiques. Parmi les sous-groupes qui ont été caractérisés fonctionnellement, certains régulent spécifiquement la construction des parois secondaires.

Afin de clarifier les relations phylogénétiques au sein des familles FT MYB et NAC chez le sorgho et de permettre de nouvelles analyses sur leur évolution par rapport à *Arabidopsis* et d'autres espèces dicotylédones et monocotylédones, nous avons agrégé les séquences protéiques des MYB et des NAC de 9 espèces différentes.

La famille NAC est une des plus grandes familles de FT chez les plantes. Une étude récente des gènes SbNAC a reporté l'existence de 14 clades NAC chez le sorgho (Sanjari *et al.*, 2019). Nos analyses complémentaires ont permis, grâce à l'intégration des séquences NAC de 9 espèces différentes, d'identifier un clade supplémentaire (P) et de séparer l'un d'entre eux en deux (L1 et L2). Seul le clade P contient seulement des gènes dicotylédones avec un grand groupe de gènes détectés chez le peuplier et l'eucalyptus, indiquant que la diversification de la famille NAC a eu lieu avant la divergence entre monocotylédones et dicotylédones. Dans la famille des NAC, le rôle de treize gènes NAC dans différentes étapes du développement des parois secondaires végétales a été validé jusqu'à présent. La plupart d'entre eux (10) font partie du clade C décrit par Sanjari *et al.* (2019), trois font partie du clade G et un fait partie du clade K.

Relativement au niveau de séparation général des espèces des différents clades chez les NAC, seul le clade F montre une séparation des gènes monocotylédones et dicotylédones remontant à un ancêtre commun lointain. Certains gènes *Brachypodium* se trouvent plus proches de gènes dicotylédones que de gènes monocotylédones, suggérant une convergence de séquence qui ne dépend pas d'une origine commune. Ce clade est également le plus éloigné des autres clades NAC identifiés. Il contient des gènes qui ont été identifiés par Sanjari *et al.* (2019), étude sur laquelle se base notre classification, comme faisant partie d'un clade dont nous n'avons pas retrouvé une grande partie des membres. Les travaux de Sanjari *et al.* (2019), montrent également que les gènes que nous avons pu retrouver et qui sont inclus dans notre clade F sont relativement éloignés phylogénétiquement du reste des NAC. L'apparition de cette sous-famille de NAC est peut-être postérieure à celle des autres clades et montre des différences de séquences importantes avec les autres NAC que nous avons identifiés ce qui expliquerait pourquoi une partie de ce clade identifié par Sanjari *et al.* (2019) ne l'est pas par nos analyses.

Comparé au clade F, les autres clades montrent une proximité plus importante des gènes de monocotylédones et de dicotylédones ainsi qu'une bonne conservation du nombre de gènes à travers les espèces. Parmi les clades contenant des gènes impliqués dans la mise en place des

parois secondaires, seul le clade K, contenant VNI2 un gène mineur impliqué dans la mise en place des parois, montre une diversification des gènes de monocotylédones par rapport aux dictotylédones (Figure II-8A).

Les groupes de gènes dont les membres des monocotylédones montrent une diversification comme celui de VNI2 sont potentiellement composés de gènes dont l'action de régulation pourrait avoir évolué pour être spécifique aux monocotylédones et liée aux différences de métabolisme et de morphologies observés entre les dicotylédones et les monocotylédones. Dans le cadre de la régulation des parois secondaires, au-delà des évidences claires de différences de morphologie des monocotylédones et des dicotylédones, des exemples de divergences de régulation existent (Rao et Dixon, 2018). En effet, des différences de patron d'expression ont été mises en évidence entre espèces pour les gènes VND qui sont chez *Arabidopsis* spécifiques des tissus vasculaires (Xiao *et al.*, 2018; Valdivia *et al.*, 2013; Yoshida *et al.*, 2013; Zhong *et al.*, 2011; Zhong et Ye, 2015b). De même, un homologue chez le riz des principaux régulateurs NAC mis en évidences chez *Arabidopsis*, OsSWN2, présente des régulations différentes de celles observées chez *Arabidopsis* (Yoshida *et al.*, 2013).

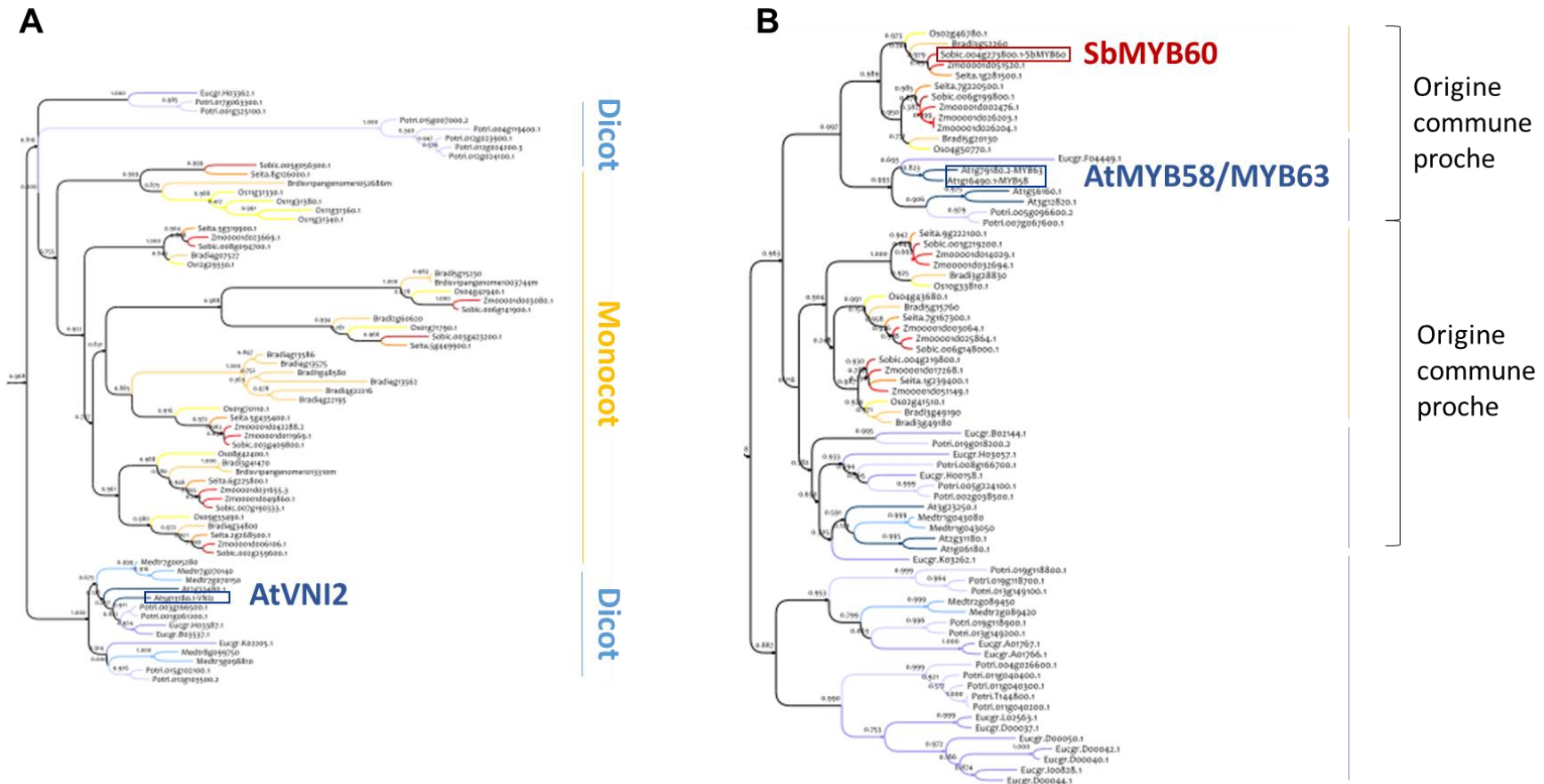


Figure II-8: Portions des arbres généraux des MYB et des NAC représentant l'organisation globale des homologues chez 9 espèces de (A) VN12 et de (B) AtMYB58 et AtMYB63. Les couleurs chaudes représentent les branches « monocotylédones » (riz, maïs, millet des oiseaux, brachypodium et sorgho) et les couleurs froides représentent les branches « dicotylédones » (luzerne, eucalyptus, peuplier et Arabidopsis). Les barres de couleur indiquent la position des groupes de gènes monocotylédones et dicotylédones. Les cadres de couleurs mettent en évidence les gènes validés fonctionnellement. Les accolades soulignent les groupes d'orthologues entre monocotylédones et dicotylédones à l'origine commune.

La famille MYB, qui est également une des plus grandes familles de FT chez les plantes, a été décrite chez plusieurs espèces comme *Arabidopsis*, le riz, le maïs, le soja et la tomate (Du *et al.*, 2012; Katiyar *et al.*, 2012; Li *et al.*, 2016; Yanhui *et al.*, 2006) et leurs rôles de contrôle de différents processus biologiques ont fait l'objet d'une revue (Dubos *et al.*, 2010). L'analyse présentée ici révèle l'existence de 135 membres de la famille MYB chez le sorgho. L'analyse phylogénétique simultanée des gènes MYB chez le sorgho et 8 autres espèces a permis l'identification de 19 clades. Comme cela a été également récemment observé avec la phylogénie des MYB de la pomme de terre (Li *et al.*, 2019), du peuplier, du riz, du maïs et de switchgrass (Zhao et Bartley, 2014) une grande cohérence existe entre la phylogénie des MYB décrite dans notre arbre multi-espèces incluant le sorgho et la phylogénie des gènes MYB décrits chez *Arabidopsis* (Dubos *et al.*, 2010; Li *et al.*, 2016).

Dix-sept FT de la famille des MYB ayant un rôle dans la mise en place des parois secondaires ont été identifiés chez *Arabidopsis*. D'après la revue de Dubos *et al.* (2010), ils appartiennent au sous-groupe S13 qui est équivalent à notre clade 4 (deux gènes et sept autres ayant des liens phylogénétiques avec ce sous-groupe), au sous-groupe S4 similaire à notre clade 15 (trois gènes), au sous-groupe S21 similaire à notre clade 12 (trois gènes) et au sous-groupe S3 similaire à notre clade 1B (deux gènes).

A l'exception du clade 3 (sous-groupe S12 selon Dubos *et al.* (2010)), qui contient exclusivement des gènes *Arabidopsis*, des séquences de dicotylédones et de monocotylédones ont été détectées de façon proche dans tous les clades ce qui suggère, comme pour la famille des NAC, une apparition de la famille des MYB avant la divergence entre monocotylédones et dicotylédones.

Cependant certains clades sont enrichis en gènes de dicotylédones. C'est le cas du clade 15 qui contient des gènes de la biosynthèse des phénylpropanoïdes (Dubos *et al.*, 2010) et notamment MYB75 qui est un régulateur important en charge du contrôle de la bascule entre la formation des parois secondaires et l'accumulation d'anthocyanes via la répression de certaines CESA et gènes de la biosynthèse des lignines ainsi que l'activation de gènes de la biosynthèse des anthocyanes (Bhargava *et al.*, 2010; Li *et al.*, 2016). La spécificité de ce gène à la lignée des dicotylédones a déjà été observée par Zhao et Bartley (2014) qui ont identifié des homologues de ce gène chez le peuplier mais pas chez les graminées. Dans le grand clade 15 que nous avons défini, AtMYB75 est effectivement situés dans une branche du sous-clade 15B où seules deux groupes de gènes monocotylédones sont également présents. Parmi eux se trouvent cependant ZmC1et ZmP11 (Allan *et al.*, 2008; Cone *et al.*, 1993) qui sont des MYB

régulateurs des anthocyanes chez le maïs. Ils ne sont pas des orthologues directs de AtMYB75 et leurs différences de rôles (l'un a un impact sur la composition des parois alors que les deux autres ne semblent avoir qu'un impact sur les anthocyanes) confirment que l'existence de branches proches mais séparées peut suggérer des divergences de fonction.

En suivant le même raisonnement, certains gènes présents dans le clade 1B, qui présente également un enrichissement en gènes de dicotylédones, sont potentiellement spécifiques des dicotylédones. Il est intéressant de noter que dans ce clade sont présents deux régulateurs principaux des parois secondaires et en particulier de la mise en place des lignines, MYB58 et MYB63 (Zhou *et al.*, 2009). Deux orthologues directs de ces gènes existent, il s'agit de SbMYB60 et de SbMYB60b. Le premier a un rôle validé dans la régulation des lignines chez le sorgho (Scully *et al.*, 2016b, 2017) et le second a été mis en évidence dans notre étude comme étant un candidat fort pour un rôle dans la mise en place des parois. Si la redondance de fonction à travers les espèces semble, dans ce cas, confirmée, une branche éloignée de celles de AtMYB58 et AtMYB63 mais appartenant au clade 1B (Figure II-8B) et contenant exclusivement des gènes dicotylédones pourrait-être, elle, impliquée dans un processus de régulation des lignines propre aux dicotylédones.

Hormis quelques expansions de groupes monocotylédones et dicotylédones qui peuvent être source de découvertes de gènes aux fonctions spécifiques de chacun de ces ordres, la bonne conservation des familles de FT MYB et NAC semble permettre l'inférence de fonction sur la base des homologies détectées. Nous nous attendons donc à ce que les informations apportées par nos analyses phylogénétiques soient utiles à la prédiction de la fonction des gènes chez des espèces chez lesquelles la caractérisation moléculaire détaillée des gènes n'est pas très développée comme c'est le cas chez le sorgho.

La connaissance des relations phylogénétiques entre les potentiels régulateurs de la mise en place des parois chez le sorgho ainsi que de l'ensemble des membres de ces familles permet de bénéficier de connaissances à coupler à des approches globales comme celles basées sur des données transcriptomiques. En effet, l'association de ces deux approches permettra d'apporter de nouveaux éléments concernant le rôle des gènes candidats identifiés par les analyses de génomique comparative.

4.2. Identification de réseaux de co-expression de gènes contribuant au développement des entre-nœuds

Une analyse des réseaux de co-expression de gènes prenant en compte différents niveaux d'entre-nœuds échantillonnés à différents stades développementaux dans des conditions différentes d'apport en eau a permis l'identification de groupes de gènes partageant les mêmes patron d'expression pendant le développement des entre-nœuds. Six groupes de gènes enrichis en GO-terms liés aux parois secondaires et aux parois en général ainsi qu'en gènes de structures liés aux parois ont été identifiés.

Quatre réseaux de gènes montrent un patron d'expression compatible avec un potentiel rôle dans l'élongation des entre-nœuds et la mise en place des parois primaires. Les deux réseaux de co-expression présentant les pics d'expression les plus précoces (les modules « turquoise » et « blue ») sont significativement enrichis en expansines et en gènes de structure des parois qui sont décrits comme impliqués dans l'élongation cellulaire (Kebrom *et al.*, 2017; McKinley *et al.*, 2016). Ils contiennent également une grande part de gènes impliqués dans la synthèse des pectines et des hémicelluloses qui sont des composés intégrés aux parois primaires. Les analyses d'enrichissement en GO-terms ont également confirmé cette spécificité avec des enrichissements en GO-terms liés aux microtubules, pectine, cellulose et hémicelluloses. Quelques membres des peroxydases et laccases sont aussi inclus dans ces réseaux de gènes. Comme mentionné par Meents *et al.* (2018), ces enzymes oxydatives peuvent avoir des rôles dans les stades précoces de la mise en place des parois. Deux autres réseaux de gènes (« lightsteelblue » et « steelblue ») qui montrent des pics d'expression légèrement retardés par rapport aux réseaux précédemment décrits sont aussi enrichis en expansines, en enzymes de relâchement des parois et en GO-terms liés aux pectines.

Deux réseaux de co-expression plus reliés à la phase de biosynthèse des parois secondaires présentent des enrichissements cohérents en GO-terms liés aux parois secondaires et en gènes de structure liés aux parois (modules « yellow » et « pink »). Ils présentent des pics d'expression synchrones avec les patrons d'accumulation de la lignine, de la cellulose et des hémicelluloses. L'appartenance de trois CESA spécifiques des parois secondaires qui ont été identifiées chez *Arabidopsis* (Taylor *et al.*, 2003) et validées chez le riz (Tanaka *et al.*, 2003) dans le module « yellow » et de deux gènes Bmr précédemment clonés chez le sorgho (dans le module « pink »: Bmr 12 - Sobic.007G047300 et Bmr2 – Sobic.004G062500) contribuent à valider la relation spécifique des gènes de ces modules avec la mise en place des parois

secondaires. Ces deux réseaux sont également significativement enrichis en laccases, peroxidases et enzymes de biosynthèse des monolignols. De façon cohérente avec l'observation d'une forte expression d'un groupe de gènes correspondant aux cyclines (impliqués dans le cycle cellulaire) dans des entre-nœuds par Kebrom *et al.* (2017), des expansines ont également été détectés dans un des deux modules liés aux parois secondaire (module « yellow »). Ces résultats soutiennent la théorie d'un maintien d'un potentiel de division cellulaire et l'élongation dans des entre-nœuds dont le développement est avancé, probablement pour permettre une croissance en diamètre.

La surexpression de CCoAOMT dans des feuilles et des tiges de sorgho a induit une réponse restreinte du transcriptome global avec seulement 1.4% des gènes impactés (Tetreault *et al.*, 2018). Parmi les 480 gènes différentiellement exprimés dans ces lignées mutées en comparaison du phénotype sauvage, 210 sont inclus dans notre analyse de co-expression. Vingt-trois pourcents d'entre eux appartient au module « yellow » et 20 % au module « turquoise », validant leur rôle potentiel dans la mise en place des parois. De plus, Scully *et al.* (2016b) ont identifié une liste de gènes co-régulés avec SbMYB60 dans des plantes sur-exprimant ce FT : la plupart de ces gènes appartiennent au module « yellow » dans notre étude, soulignant de nouveau le rôle potentiel des gènes de ce module dans la mise en place des parois secondaires.

Ces résultats renforcent les travaux pionniers réalisés par le groupe de Mullet (Kebrom *et al.*, 2017; McKinley *et al.*, 2016) dans la description des réseaux de co-expression impliqués dans le développement de l'entre-nœud. Cela ouvre aussi la voie à l'identification de gènes qui ne correspondent pas à des gènes de structure des parois déjà connus. Il est particulièrement intéressant de noter que parmi les différents modules identifiés et liés aux paroi, à l'exception du module « lightsteelblue » qui contient uniquement 9 gènes, entre 86 et 97% des gènes ne correspondent pas à des gènes déjà identifiés comme impliqués dans l'établissement des parois (655 gènes liés aux parois ont été identifiés à travers une revue bibliographique), laissant la place à une caractérisation de ces potentiels nouveaux acteurs de la biosynthèse et de la régulation des parois.

4.3.Évaluation des liens d'homologies entre les NAC et les MYB par l'analyse des réseaux de co-expression

En intégrant les données de phylogénie avec les analyses de réseau de co-expression, un niveau triple d'information est disponible pour tous les MYB et NAC présents dans le génome du sorgho (i.e. phylogénie, profil d'expression et annotation des gènes présents dans le « sous-réseau » de gènes le plus co-exprimés avec eux). Globalement, les analyses entreprises sur les MYB et les NAC ont révélé que pour la plupart des sous-clades contenant des FT avec un rôle décrit dans la mise en place des parois secondaires, les gènes homologues chez le sorgho présentent des patrons d'expression et des relations de co-expression cohérents avec ces mêmes rôles. En effet, du premier niveau des régulateurs principaux jusqu'aux régulateurs aval pour lesquels des évidences fonctionnelles existent chez *Arabidopsis* et quelques autres espèces, nous avons toujours pu identifier des FTs MYB et NAC chez le sorgho dont le même rôle est suspecté. Une synthèse de ces résultats est disponible en Tableau II-6.

Les gènes *Arabidopsis* NST1 et NST3/SND1 sont les régulateurs principaux de la mise en place des parois secondaires chez *Arabidopsis*. Les deux homologues identifiés chez le sorgho appartiennent au module « Yellow » qui est très relié aux parois secondaires et les deux gènes présentent des sous-réseaux fortement enrichis en gènes et GO-terms liés aux parois secondaires. Ces observations sont cohérentes avec les résultats chez le maïs, le riz et plus récemment chez switchgrass (Rao *et al.*, 2019). En plus des régulateurs principaux de la mise en place des parois secondaires dans les fibres (NST), des FT spécifiques des vaisseaux et clés du développement des parois secondaires ont également été identifiés chez *Arabidopsis* (VND). Nous avons montré qu'un orthologue de VND7 (Sobic.007G003000 / SbVND7a) qui appartient également au module « yellow », présente un sous-réseau enrichi en gènes des parois secondaires et qu'il participe probablement à la régulation de 4 MYB (SbMYB83, SbMYB61a, SbMYB92a et SbMYB43) (il fait partie de leur sous-réseau de forte co-expression) dont deux sont orthologues de régulateurs en aval identifiés chez d'autres espèces. Le fait que nous n'ayons pas observé d'expression des homologues d'AtVND1/2/3 est cohérent avec les récents travaux de Tan *et al.* (2018) qui ont suggéré que ces gènes ont un rôle de régulateur pendant le développement des jeunes plantules chez *Arabidopsis* puisque nos échantillons ont été prélevés plus tardivement que ces stades juvéniles. Ces résultats sont également cohérents avec la forte expression de ces deux gènes dans les racines et les tiges

d'échantillons provenant de plantules (Base de données MOROKOSHI : <http://sorghum.riken.jp/morokoshi/Home.html>).

Le deuxième niveau des régulateurs de la mise en place des parois secondaires chez *Arabidopsis* implique les gènes redondants AtMYB46 et 83. Le seul gène orthologue de ces deux gènes chez le sorgho appartient au module « yellow », qui est associé aux parois secondaires, et son sous-réseau est également fortement enrichi en GO-terms liés aux parois. De plus, il appartient aussi au sous-réseau de SbMYB61a, un homologue de AtMYB61, un régulateur aval des parois secondaires. Parmi les FT du groupe de régulation aval validés chez *Arabidopsis*, nous avons montré, entre autres, des évidences claires pour la conservation du rôle de AtMYB103 chez le sorgho. En effet, le seul orthologue existant chez le sorgho a un patron d'expression (appartenant au module « pink ») et des gènes co-exprimés fortement en faveur d'un rôle dans la régulation des parois secondaires comme le sous-entend son lien phylogénétique.

Au-delà de fournir de bons indices sur la conservation fonctionnelle entre espèces au niveau du gène (comme observé par Rao *et al.* (2019) entre switchgrass et *Arabidopsis*), l'intégration de la phylogénie et de l'analyse de co-expression de gènes permet d'identifier des clades MYB et NAC présentant des fonctions différenciées. En effet, les gènes appartenant au clade NAC C et aux clades MYB 1B, 4, 11 et 12 sont potentiellement capables de contribuer à la mise en place des parois primaires et secondaires tandis que les gènes NAC du clade G et les MYB du clade 15 rassemblent des gènes qui seraient plus spécifiques des parois secondaires.

Le clade NAC C contient 13 gènes de sorgho, dont quatre correspondent aux orthologues précédemment mentionnés de NST1/2/3 et de VND7 qui sont liés aux parois secondaires. Il contient également deux gènes orthologues de AtVND4/5/6 (Sobic.010G002900, Sobic.004G302400) qui sont spécifiquement exprimés dans les premières étapes de l'établissement des parois (élongation et synthèse de la cellulose, des hémicelluloses et de la pectine).

La même chose a été observée pour le clade MYB 1B qui inclut SbMYB60b qui est lié aux parois secondaires (comme montré aussi pour son paralogue SbMYB60 par Scully *et al.* (2016b)) et SbMYB7 qui est plutôt lié aux parois primaires (membre du module « blue » avec des expansines, des pectates et pectine lyase et plusieurs XTH). Les gènes des clades MYB 4, 11 et 12 présentent les mêmes variations de patron d'expression et d'enrichissement de sous-réseaux.

Au contraire, des clades qui semblent spécifiques des parois secondaires ont été identifiés. Le clade NAC G, qui contient des homologues et des gènes proches de AtNAC075, contient exclusivement des gènes avec des preuves d'un rôle dans la mise en place des parois secondaires. De façon identique, le clade MYB 15, qui contient par exemple les homologues de AtMYB4 (qui est validé chez *Arabidopsis* comme un répresseur des lignines), rassemble des gènes plutôt liés aux parois secondaires.

En plus des FT MYB et NAC, l'analyse des réseaux de co-expression a permis d'identifier un potentiel rôle dans la mise en place des parois pour les homologues chez le sorgho des gènes WRKY12 et KNAT7 qui sont des gènes clés de ce processus chez *Arabidopsis*. En effet, ces deux FT font partie des « hub-genes » du module « pink » qui est un module très lié aux parois secondaires. Pour WRKY12, ces résultats sont cohérents avec ceux précédemment reportés chez les graminées (Gallego-Giraldo *et al.*, 2016; Li *et al.*, 2012; Wang *et al.*, 2010) et avec les récents travaux chez switchgrass par Rao *et al.* (2019) qui ont souligné un potentiel rôle de répresseur de ce gène sur la biosynthèse des lignines.

Globalement, nous avons observé une relativement bonne capacité de prédiction des fonctions des homologues « sorgho » des MYB, NAC et de quelques autres FT précédemment validés chez *Arabidopsis* et d'autres espèces. Cependant, nous avons rapporté quelques divergences entre le sorgho et la plante modèle *Arabidopsis thaliana*. La première divergence entre le sorgho et *Arabidopsis* est l'absence de mise en évidence des homologues de AtMYB69, un potentiel acteur de la régulation des parois secondaires médiée par SND1 et NST1 (Zhong *et al.*, 2008). Ce résultat est cohérent avec ceux rapportés chez switchgrass où une absence d'expression des deux homologues de AtMYB69 avait également été observé (Zhao et Bartley, 2014). Dans ce cas l'hypothèse de la régulation ontogénique proposé pour VND1/2/3 par Tan *et al.* (2018) n'est pas valable car des échantillons de plantules ont été pris en compte dans les analyses transcriptomiques chez switchgrass. Une autre divergence avec *Arabidopsis* est le rôle potentiellement mineur des homologues du sorgho de VND7 en comparaison avec les autres régulateurs principaux. Même s'il a été montré chez *Arabidopsis* que VND7 régule positivement directement et/ou indirectement de nombreux gènes impliqués dans une large gamme de processus de la différenciation des vaisseaux du xylème (Yamaguchi *et al.*, 2011), nos résultats chez le sorgho sont moins clairs. En effet, seul des enrichissements en gènes impliqués dans la biosynthèse des cires et des pectines ont été détectés dans le sous-réseau fortement co-exprimés avec SbVND7a.

Toutefois, avant de rejeter un rôle de ces FT MYB et NAC dans la mise en place des parois de la tige chez les graminées des analyses complémentaires au niveau du tissu ou des groupes de cellules sont nécessaires.

Comme montré ci-dessus il est assez évident que l'inférence de fonctions de gènes chez le sorgho peut être ancrée sur les résultats rapportés à partir *Arabidopsis* en faisant attention d'identifier quels homologues présentent les mêmes caractéristiques à travers l'analyse d'informations transcriptomique.

Nous avons identifié 15 nouveaux MYB et cinq nouveaux NAC pour lesquels un rôle potentiel dans la mise en place des parois peut être proposé. Ces FT appartiennent aux six modules de co-expression précédemment identifiés par leur lien avec les parois primaires et secondaires et ne sont pas des homologues proches des gènes impliqués dans la régulation des parois secondaires déjà caractérisés chez *Arabidopsis*. La moitié de ces gènes n'appartiennent pas aux clades précédemment identifiés comme rassemblant des gènes ayant un rôle dans la régulation des parois. Il est intéressant de souligner que pratiquement tous (ce qui inclut deux groupes de paralogues) appartiennent à des clades qui semblent avoir évolué séparément des dicotylédones. Ces résultats suggèrent des spécificités dans la régulation de la mise en place des parois primaires et secondaires chez le sorgho et les graminées par rapport aux dicotylédones.

Même si une analyse approfondie de ces nouveaux FT liés aux parois reste à réaliser, il est intéressant de souligner que les deux homologues uniques (LOC_Os05g35500 et LOC_Os12g07640) des 4 gènes de sorgho SbMYB92a (Sobic.009g148100), SbMYB92, (Sobic.008g055800), SbMYB92b (Sobic.005g062000) et SbMYB92c (Sobic.008g055700) ont également été identifiés par Hirano *et al.* (2013a) grâce à une analyse de réseaux de co-expression similaire à la nôtre.

En prenant en compte les nouveaux éléments obtenus dans notre étude, nous proposons un schéma de la régulation de la mise en place des parois secondaires chez le sorgho (Figure II-9).

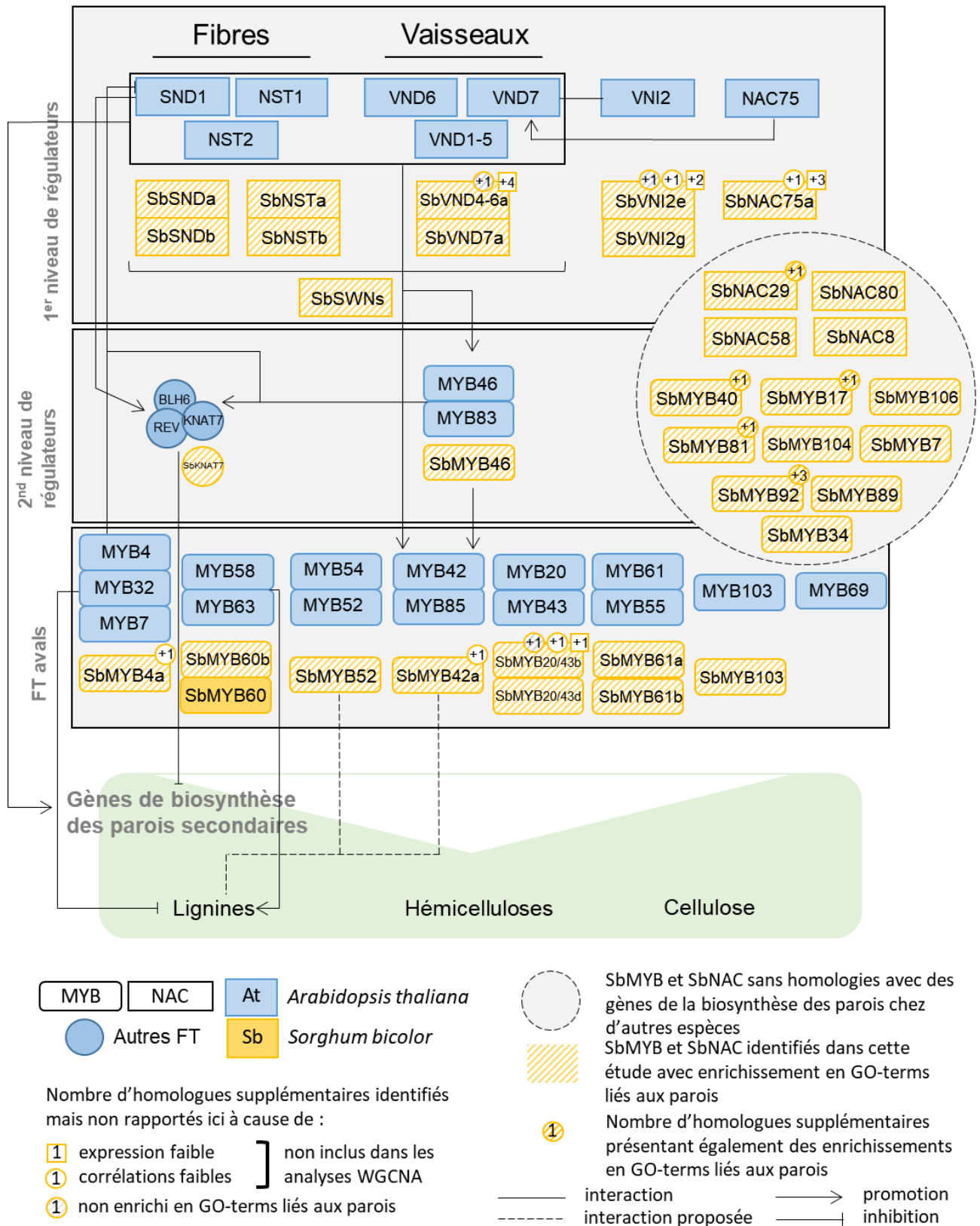


Figure II-9 : Facteurs de transcription MYB et NAC et autres régulateurs des parois secondaires validés fonctionnellement chez *Arabidopsis* et leurs homologues chez le sorgho identifiés dans cette étude par analyse de réseaux de co-expression de gènes.

En comparaison Kebrom *et al.* (2017) qui ont rapporté une co-régulation générale des MYB et des NAC au sein des 4 entre-nœuds apicaux de sorgho (avec une expression augmentée dans l'entre-nœud le plus âgé), nous avons observé une plus grande diversité de patrons d'expression, avons mis en lumière les points communs et les divergences avec *Arabidopsis* et permis d'affiner la liste des FT les plus intéressants dans le contexte de la mise en place des parois pour des expérimentations de validation.

Chez le sorgho, des FT NAC (mais aucun FT MYB) ont été détectés lors d'études sur la diversité nucléotidique en lien avec la composition de la biomasse. Un premier lien a été rapporté entre SbNSTb (Sobic.010G022400) et le rendement de saccharification par une analyse GWAS (Wang *et al.*, 2013a). Nos résultats sont en accord avec une implication potentielle de ce gène dans la mise en place des parois car ce gène appartient au module « yellow » qui est lié à la mise en place des parois secondaires et présente un sous-réseau enrichi en gènes liés aux parois qui sont principalement impliqués dans la biosynthèse de la cellulose et des hémicelluloses. Une deuxième étude concerne le gène NAC sous-jacent au locus D (Sobic.006G147400, SbNAC074) qui contrôle la jutosité de la tige. Les résultats mis en évidence par Casto *et al.* (2018) et Xia *et al.* (2018) suggèrent que ce gène est potentiellement responsable du contrôle de la mort cellulaire et module la proportion d'aérenchyme dans la partie centrale des tiges de sorgho (Casto *et al.*, 2018; Xia *et al.*, 2018). Sur la base de nos résultats, ce gène appartient au clade B mais il n'a pas été pris en compte dans nos analyses à cause de la faible corrélation d'expression qu'il présente avec les autres gènes. Ce gène pourrait agir à un niveau différent de la mise en place des parois comme dans les mécanismes d'accumulation en sucres solubles et dans la structuration du parenchyme.

4.4. Gènes associés à la qualité de la biomasse en lien avec les FT NAC et MYB et les réseaux de co-expression liés à la mise en place des parois

L'intégration des données de génétique d'association liées à la qualité de la biomasse, les connaissances des homologues chez le sorgho des principaux régulateurs des parois secondaires et des réseaux de co-expression reliés aux parois a permis l'accumulation d'informations nécessaires à l'analyse des gènes candidats potentiellement acteurs de la mise en place des parois chez le sorgho.

Parmi les enzymes connues pour leur implication dans la biosynthèse des parois (liste issue de la littérature), 5 sont également présentes dans les réseaux de co-expression construits à partir de données transcriptomiques acquises pendant le développement d'entre-nœuds de sorgho. Elles sont toutes des enzymes impliquées dans la biosynthèse des polysaccharides. L'une d'entre elle (Sobic.004G051300) est une galactosyltransferase responsable de transfert de UDP-glucose à un accepteur indéterminé (protéine, polysaccharides etc.) et 4 d'entre elles sont impliquées dans la biosynthèse et/ou dans le remodelage des hémicelluloses (Tableau II-7).

En effet, les deux xyloglucanes galactosyltransferase du module "yellow" font partie de la famille des GT47 (Xu *et al.*, 2018) qui participent à la biosynthèse des xylan et des xyloglucanes. En particulier IRX10 (AT1G27440) est l'homologue le plus proche de Sobic.009G220200 et est essentiel à la formation de la structure des xylandes dans les parois secondaires d'*Arabidopsis* (Brown *et al.*, 2009; Wu *et al.*, 2009; Jensen *et al.*, 2014; Urbanowicz *et al.*, 2014, Rennie et Scheller, 2014; Zhang *et al.*, 2014). Chez les monocotylédones, les hémicelluloses principales de la paroi secondaire sont les xylandes tandis que les xyloglucanes y sont très mineures (Scheller et Ulvskov, 2010). Les xylandes sont également présentes dans les parois primaires en quantité inférieure aux parois secondaires et les xyloglucanes y représentent un faible pourcentage (Scheller et Ulvskov, 2010). Ces deux enzymes sont associées à la variabilité des parois dans les données de génétique d'association analysée. Elles font également partie du module "yellow" qui est associé aux parois secondaires ainsi que du sous-réseau de SbMYB17 qui est un très fort candidat de la régulation des parois chez le sorgho mis en évidence par les analyses de réseaux de co-expression. Enfin, elles font partie des 25 gènes les plus centraux ("hub genes") du module yellow (en position 13 et 24) (Tableau II-7). Ces éléments suggèrent que ces deux enzymes sont centrales à la mise en place des parois chez le sorgho et potentiellement impliquées spécifiquement dans l'établissement des parois secondaires.

De plus, deux autres enzymes impliquées dans la biosynthèse des hémicelluloses sont associés aux méta-traités de composition de la biomasse détectés en génétique d'association et font partie des modules de co-expression associés aux parois identifiés au cours des analyses de réseaux de gènes. Sobic.004G308600 est une xyloglucan fucosyltransferase homologue du gène *Arabidopsis* FT1 (Chiniquy *et al.*, 2012) et associée aux métra-traités de contenu en parois et de sucres solubles (Tableau II-7). Elle a également été détectée dans une étude

transcriptomique de dynamique d'expression des gènes sur des entre-nœuds de sorgho en développement (McKinley *et al.*, 2016). Son appartenance au module de co-expression "turquoise" qui est relié à la mise en place des parois primaires suggère qu'elle est impliquée dans la biosynthèse des hémicelluloses de la paroi primaire chez le sorgho.

Sobic.002G194500, une XTH dont le rôle putatif est le remodelage des xyloglucanes est associée aux traits de contenu en paroi et de digestibilité dans les analyses de génétique d'association étudiées. Les liaisons entre les polysaccharides des parois étant garantes de leur résistance et les XTH étant associées au relâchement des parois, la variabilité de l'activité de cette XTH pourrait donc effectivement entraîner une modification de l'accessibilité des sucres des parois et de la digestibilité de la biomasse (Cosgrove, 2016). Cette enzyme fait également partie du module de co-expression "blue" associé aux parois primaires et est intégrée au sous-réseau de régulation de SbMYB7, un bon candidat (catégorie "moyennement enrichi") de la régulation des parois chez le sorgho mis en évidence par l'analyse des réseaux de co-expression (Tableau II-7). Ces éléments suggèrent que cette enzyme pourrait, elle aussi, être impliquée dans la mise en place des parois chez le sorgho.

En plus des quatre enzymes mentionnées ci-dessus, Sobic.008G171600, est associé au métra-trait de contenu en paroi et intégré dans le module de co-expression "pink" qui est relié aux parois secondaires chez le sorgho. Ce gène est homologue à TT12, une protéine de type MATE (Multidrug and Toxic compound Extrusion) impliquée dans l'accumulation de proanthocyanides dans le manteau des graines d'*Arabidopsis* (Marinova *et al.*, 2007) et capable de transporter des monomères glucosylés de flavanols dans la vacuole de cellules de graines d'*Arabidopsis* (Zhao *et al.*, 2009). Elle est également intégrée au réseau de co-régulation de SbMYB52/54 un homologue de AtMYB52 et AtMYB54 qui sont des régulateurs principaux des parois secondaires chez *Arabidopsis* (Zhang *et al.*, 2018) et un bon candidat de la régulation de ce processus chez le sorgho (Tableau II-7). Ainsi, malgré le fait que TT12 soit impliqué dans un processus localisé dans les grains, son homologue chez le sorgho est exprimé dans les entre-nœuds en développement. La présence de différents flavonoïdes dans les parois (Ermeidan *et al.*, 2012) et l'interconnexion de leur voie de biosynthèse avec celles des lignines (Lam *et al.*, 2017) est compatible avec un impact de la variabilité de cette protéine sur les caractères de contenu en paroi mesurés par l'étude de génétique d'association analysée dans ce travail.

Les autres protéines pour lesquelles aucun lien avec des listes d'enzymes de la biosynthèse des parois n'est connu mais qui font partie des groupes de co-expression liés aux parois

constituent également des candidats pour un rôle dans la mise en place des parois chez le sorgho, en représentant des gènes « nouveaux » par rapport à ce qui est décrit dans la littérature. En particulier, trois gènes remplissent ces conditions, qui font partie des « hub gene » du module de co-expression « turquoise » (Tableau II-7), cependant, des études complémentaires sont nécessaires pour conclure sur leur rôle.

Enfin, les deux FT associés au contenu en parois sont les seuls détectés dans les essais de génétique d'association analysés dans ce travail. L'un n'est pas inclus dans les réseaux de co-expression tandis que l'autre, un homologue de VND4, un régulateur des parois chez *Arabidopsis* (Zhou *et al.*, 2014), n'est pas mis en évidence par l'étude de ces réseaux comme un bon candidat de ce processus chez le sorgho. L'absence des FT sous-jacents aux zones d'intérêt évaluées ici est cohérente avec la faible détection des gènes MYB et NAC en génétique d'association chez le sorgho. En effet, seuls les homologues de NST1 et de NAC074, deux FT NAC, ont été reliés à des caractères de composition de la biomasse chez le sorgho (Burks *et al.*, 2015; Wang *et al.*, 2013b; Xia *et al.*, 2018; Zhang *et al.*, 2018). Ce manque de détection peut être causé par différents points. Ainsi, les principaux FT régulateurs des parois secondaires étant cruciaux dans ce processus déterminant pour la fitness des individus, leurs séquences pourraient être très peu variables parmi les accessions disponibles pour mener des études d'association et ainsi demander des panels de diversité extrêmement larges permettant leur détection. Les conditions environnementales dans lesquelles les essais de génétique d'association ont été menés n'ont peut-être pas non plus permis l'expression de leur variabilité. L'absence de détection de ces gènes peut également être liée aux mesures phénotypiques réalisées au cours des essais de génétique d'association. En effet, l'action des FT sur le phénotype est robuste et régulièrement renforcée par des boucles de régulation qui compensent leur potentielle variabilité d'expression (Payne et Wagner, 2015). Ainsi, les mesures phénotypiques peuvent ne pas permettre de détecter la variabilité phénotypique induite par leur variabilité génétique. L'effet phénotypique induit peut également être subtil et, la quantification de la composition de la biomasse utilisée reposant ici sur des mesures non spécifiques à des composés précis de la tige, des mesures plus précises pourraient être nécessaires pour la mise en évidence d'association avec les FT régulateurs des parois secondaires.

5. Conclusion

Le travail rapporté dans ce chapitre a permis l'identification de différents gènes potentiellement impliqués dans la mise en place des parois chez le sorgho grâce à la multiplication des approches (génomique comparative, réseaux de co-expression et génétique d'association). Les FT et les gènes de structure les plus intéressants sont reportés dans le Tableau II-7 et dans la Figure II-9. L'identification de gènes candidats est une des étapes permettant d'atteindre une meilleure compréhension des mécanismes du dépôt et de la régulation des parois secondaires. Ainsi, afin d'explorer les fonctions des gènes candidats nous avons mis en place des essais de transformation génétique dans l'objectif de valider fonctionnellement le rôle de ces gènes.

**Chapitre III : Utilisation de la transformation
génétique pour la validation fonctionnelle de gènes
impliqués dans la mise en place des parois
secondaires**

1. Introduction

Les analyses décrites dans le chapitre II ont permis de mettre en évidence des gènes candidats potentiellement impliqués dans la mise en place des parois secondaires chez le sorgho. Afin d'être en mesure de valider leur implication dans ce processus et d'éclaircir leur rôle, nous avons souhaité utiliser une approche de génétique inverse. La génétique inverse repose sur l'obtention de plantes mutantes aux loci d'intérêts dont le phénotype, une fois caractérisé, informe sur le rôle des gènes ciblés par les mutations. Plusieurs stratégies d'obtention de mutants existent, elles se divisent en deux catégories selon que les mutations sont aléatoires ou ciblées. Les mutations induites aléatoirement sont généralement utilisées pour la recherche haut-débit de gènes d'intérêt (génétique directe) mais peuvent également permettre de générer des banques de mutants dans lesquelles sélectionner les lignées qui portent des mutations aux loci d'intérêt. Les méthodes d'obtention peuvent être chimiques (EMS), physiques (rayons UV, α et γ) ou biologiques (T-DNA, transposons). Une banque de mutants générés aléatoirement au travers de l'utilisation d'un mutagène chimique (EMS) est disponible chez le sorgho (Jiao *et al.*, 2016). Quatre-vingt-quinze pourcents des gènes du génome du sorgho sont touchés par des mutations parmi les 256 lignées qui ont été entièrement séquencées (30X). Il est donc potentiellement possible de sélectionner une lignée mutante pour chacun des gènes d'intérêt ciblés. Cependant, l'impact des mutations induites n'est pas contrôlée et la densité de mutation est forte dans le génome de ces plantes. L'effet de la mutation du gène ciblé peut donc être nul en cas de mutation synonyme et/ou difficile à exploiter du fait du grand nombre de gènes potentiellement mutés. L'absence de contrôle sur le locus précis de mutation est un des défauts principaux qui se retrouvent dans toutes les méthodes aléatoires de mutations chimique et physique et dans une certaine mesure également chez les plantes dont les mutations sont induites par l'insertion non contrôlée de T-DNA et de transposon. Nous avons donc choisi de travailler avec des mutants obtenus de façon ciblée par transformation génétique. Nous avons également développé une méthode de transformation transitoire sur protoplastes permettant de diminuer le temps d'obtention de résultats.

1.1.Principe et intérêt de la transformation génétique

La transformation génétique ciblée permet l'introduction d'un matériel génétique donné dans du matériel végétal. L'objectif du transfert de matériel génétique, appelé transgène et qui peut être de nature variable, est d'apporter à la plante transformée l'information et la fonctionnalité qu'il porte. Le transgène peut être une version modifiée d'un gène de l'espèce transformée

(que ce soit un allèle ou des séquences régulatrices différentes), un gène issu d'un autre organisme ou encore un complexe qui a vocation à modifier la structure ou l'expression d'un gène de la plante transformée. Il existe deux types de transformations génétiques, l'une stable dont les effets sont durables et transmissibles à la génération suivante et qui fait l'objet de cette première partie d'introduction ; et l'autre, transitoire, qui fait l'objet de la dernière partie d'introduction et dont l'objectif est d'observer rapidement des effets phénotypiques, en particulier des effets transcriptomiques, et qui n'a pas forcément vocation à mener à la constitution de lignées de plantes mutées.

La première insertion stable de matériel exogène a été décrite en 1984 chez le tabac (Block *et al.*, 1984; Horsch *et al.*, 2016; Paszkowski *et al.*, 1984) et a depuis essaimé dans 35 familles de plantes (Keshavareddy *et al.*, 2018). Face aux défis de la recherche fondamentale et appliquée, la transformation génétique stable s'est avérée être un outil indispensable. Son utilisation en recherche fondamentale a notamment permis la caractérisation et l'identification de nombreux gènes. En effet, la transformation génétique permet d'étudier l'effet de modifications précises sur le phénotype des organismes puis d'en déduire des fonctions, interactions et positionnement dans les cascades de régulation des processus biologiques.

La recherche appliquée a également pu bénéficier des apports de la recherche fondamentale et mettre en place ses propres approches de transformation génétique ciblée. En considérant les processus de sélection classique qui s'étalent sur de nombreuses années voire dizaines d'années, les techniques de transformation génétique sont particulièrement attractives car elles accélèrent les découvertes et ont le potentiel de mener à des variétés commercialisables en quelques années (Altpeter *et al.*, 2016; Sanagala *et al.*, 2017). De nombreux traits ont été et pourraient être la cible de cette approche, en particulier les résistances aux pathogènes, aux herbicides et aux insectes ravageurs, la tolérance aux conditions climatiques, une conversion énergétique plus efficace à travers la modification de la photosynthèse, la fixation de l'azote de l'air ou encore les caractéristiques nutritives, de digestibilité, de phytoremédiation (Nicholl, 2002; Sant, 2011; Slater *et al.*, 2003).

Au cours du travail que rapporte cette thèse, l'ambition était d'utiliser la transformation génétique comme un outil exploratoire des processus biologiques visant à mettre en évidence et valider le rôle de gènes impliqués dans la mise en place des parois.

1.2.Principe de transfert de l'ADN dans le matériel végétal

Les trois prérequis pour la transformation génétique stable sont l'existence d'un tissu « compétent » pour recevoir le matériel génétique et capable de se régénérer en une plante entière, une méthode d'introduction du matériel génétique dans ce tissu et une procédure pour sélectionner et régénérer des plantes avec un taux de succès satisfaisant.

Un des premiers défis de la transformation est donc, en amont de la méthode de transfert d'ADN, la régénération des tissus qui ont subi cette transformation. En effet, en utilisant les capacités de totipotence et de plasticité des cellules végétales, il est possible de régénérer une plante entière à partir d'une cellule transformée, c'est l'objectif de la culture *in vitro* de tissus végétaux dans des conditions contrôlées et aseptiques.

Les méthodes de régénération utilisent majoritairement des protoplastes (cellules individualisées), des embryons (cellule somatique) ou l'organogénèse directe à partir de fragments de tiges et de feuilles (Bhojwani et Razdan, 1996). Si ces trois méthodes aboutissent toutes à la formation d'une plantule, l'embryogénèse somatique à partir d'explants implique le développement d'un embryon qui germe et se développe en une jeune plante (Figure III-1A et B) alors que l'organogénèse induit la formation d'organes (tige et racine) à partir de tissus végétatifs (Figure III-1C). L'utilisation de protoplastes et d'embryons permet l'embryogénèse somatique indirecte avec des cals issus de cultures cellulaires elle-même issues de tige et des cals issus d'embryons. Ces deux premières méthodes (Figure III-1A et B) dépendent beaucoup du génotype et de l'explant utilisé ce qui signifie que le succès d'un protocole n'est pas assuré à travers tous les idéotypes et cultivars d'une même espèce. Ce sont des méthodes très longues qui demandent énormément de travail à la fois pour leur optimisation et pour leur mise en œuvre en routine (Christou, 1992). L'organogénèse directe (Figure III-1C) à partir de fragment végétaux est moins sensible au génotype (Harshavardhan *et al.*, 2002; Zhong *et al.*, 1998) mais peut aisément conduire à la formation de plantes chimériques (constituées d'un assemblage de cellules qui n'ont pas le même patrimoine génétique) (Sant, 2011). Les variations somaclonales sont également observées en culture *in vitro* de cals et en cultures de suspensions cellulaires. Il s'agit de réarrangements chromosomiques qui affectent le nombre de chromosomes et leur structure induisant de grands changements phénotypiques (Hansen et Wright, 1999; Rakoczy-Trojanowska, 2002). Même si de nombreuses techniques de culture *in vitro* peuvent être utilisées, il n'existe pas de méthode universelle applicable chez toutes les espèces et il est difficile de contrôler tous les

paramètres de sa mise en œuvre (Sant, 2011). Dans certains cas idéaux, la culture *in vitro* n'est pas requise car la transformation peut être réalisée sur les organes floraux qui produisent des grains matures transformés. Cette technique est massivement utilisée chez *Arabidopsis thaliana* (Clough et Bent, 1998; Koetle *et al.*, 2015) et chez d'autres plantes, incluant des monocotylédones, mais de façon non standard et difficilement reproductible (Mu *et al.*, 2012; Ratanasut *et al.*, 2017; Zale *et al.*, 2009).

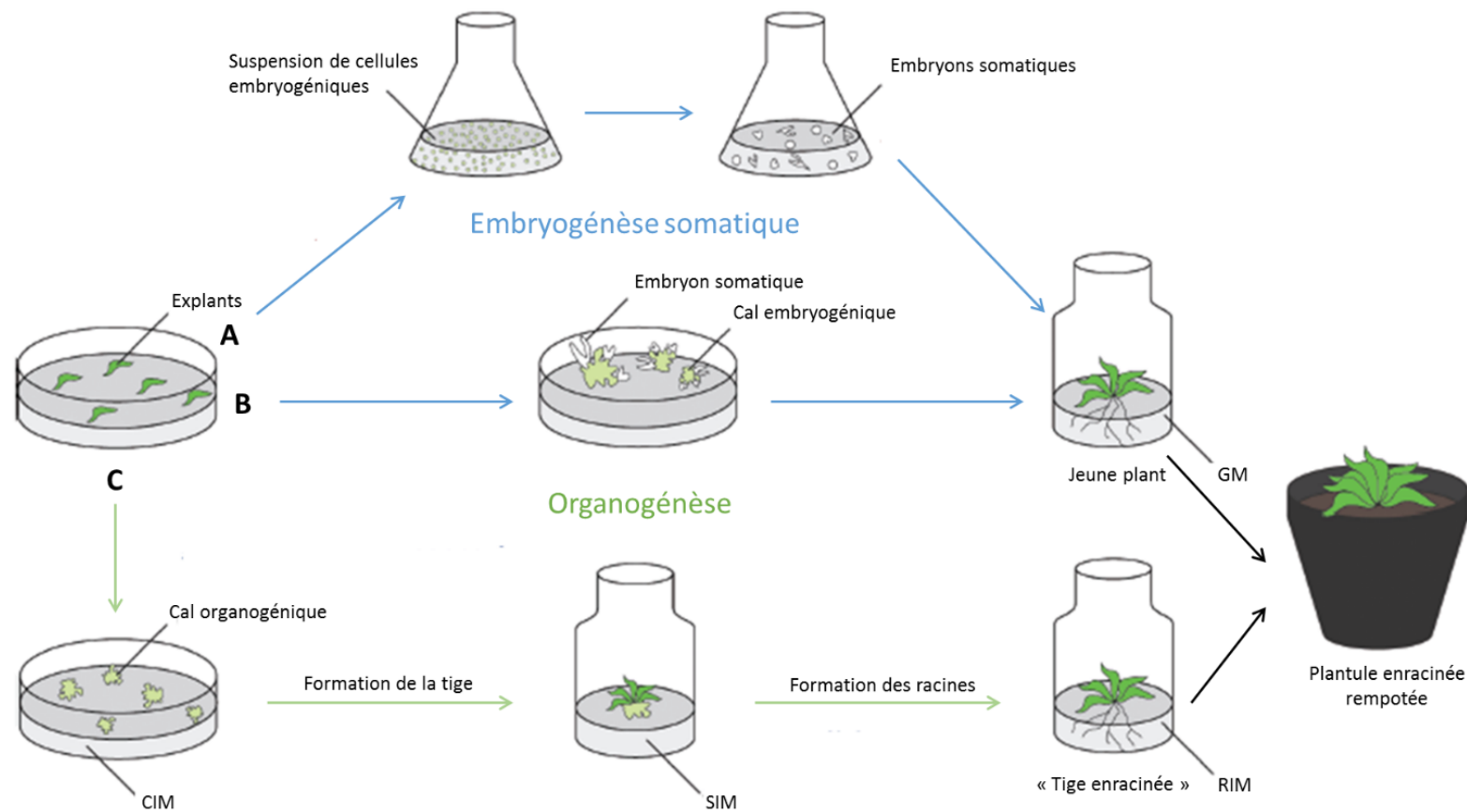


Figure III-1 : Embryogénèse somatique et régénération de tige. (A) Embryogénèse somatique avec culture de cellules embryogéniques. (B) Embryogénèse somatique avec culture de cals embryogéniques. (C) Régénération de tiges avec cals organogéniques. CIM=Callus Initiation Medium, SIM=Shoot Initiation Medium, RIM=Root Initiation Medium, GM=Germination Medium (d'après Anami *et al.*, 2013)

Lorsque sont développés des systèmes efficaces de régénération de plantes entières, la méthode de transfert d'ADN peut-être un autre goulot d'étranglement des protocoles. Deux classes de systèmes de transfert génétiques existent : les méthodes biologiques et les méthodes non biologiques qui incluent les méthodes chimiques et physiques. Chez les plantes, les deux principales méthodes de transformation génétique stable reposent sur l'utilisation d'un vecteur bactérien, *Agrobacterium tumefaciens* (méthode biologique), et sur le bombardement direct de matériel génétique par biolistique (méthode physique). Les autres méthodes que sont par exemple l'électroporation ou la sonication, sont mineures (Anami *et al.*, 2013; Keshavareddy *et al.*, 2018).

La transformation avec le vecteur *Agrobacterium* repose sur l'aptitude naturelle de cette bactérie à transférer son matériel génétique en utilisant sa propre machinerie cellulaire et celle de la plante hôte (Gelvin, 2003). Pendant l'infection, *Agrobacterium* réplique un fragment de son ADN qu'on appelle le T-DNA (Transferred DNA) et le transfère dans les cellules de la plante hôte où il s'intègre dans le génome. Normalement, les T-DNA transférés contiennent des gènes qui entraînent une prolifération cellulaire (biosynthèse de phytohormones) et permettent la sécrétion par la plante de dérivés d'acides aminés consommés par la bactérie. Même si *Agrobacterium* infecte naturellement en majorité des plantes dicotylédones (De Cleene et De Ley, 1976), elle peut transformer d'autres espèces en conditions de laboratoire et est devenue un vecteur de choix de la transformation génétique chez la plupart des espèces (Tzfira et Citovsky, 2006). Le développement des techniques de transformation génétique est passé par le remplacement des gènes originellement présents dans le T-DNA par d'autres gènes comme des gènes de sélection et des gènes d'intérêt (GI). La construction de cette séquence d'intérêt est répliquée dans un système exogène bactérien (*E. coli*) puis transférée dans des bactéries *Agrobacterium* qui conservent leur matériel génétique d'assistance au transfert et de virulence natifs (*vir*) qui permettent l'intégration du T-DNA d'intérêt (Gelvin, 2003). Une co-culture des bactéries transformées et des tissus végétaux est la méthode principale pour initier l'infection et le transfert génétique.

La deuxième méthode très utilisée en transformation génétique chez les plantes est la biolistique. Il s'agit d'une méthode directe physique décrite pour la première fois en 1987 (Sanford *et al.*, 1987) qui visait à fournir une alternative aux autres méthodes de transformation en particulier chez les céréales, particulièrement récalcitrantes (Sant, 2011; Smith et Hood, 1995; Sood *et al.*, 2011). Elle est toujours très utilisée pour sa meilleure efficacité notamment chez les graminées mais aussi car elle permet le transfert de nombreux

gènes simultanément (jusqu'à 14 chez le riz (Chen *et al.*, 1998)), d'éviter les problèmes de contaminations bactériennes incontrôlées et enfin en raison de la facilité de construction des vecteurs et de leur administration (Dai *et al.*, 2001; Kaur et Sah., 2014). Le bombardement peut, comme la transformation par *Agrobacterium*, être réalisé sur plante entière (Hamada *et al.*, 2017) mais il est le plus souvent appliqué à des explants. La biolistique repose sur l'utilisation de billes d'or comme vecteur d'ADN (Kaur et Sah., 2014). Un mélange de billes et de plasmides portant le transgène est propulsé à grande vitesse par un canon à particules sur le matériel végétal. Les billes bombardées perforent les tissus végétaux et permettent au transgène d'accéder à l'intérieur de la cellule où il s'intègre de façon non contrôlée dans l'ADN.

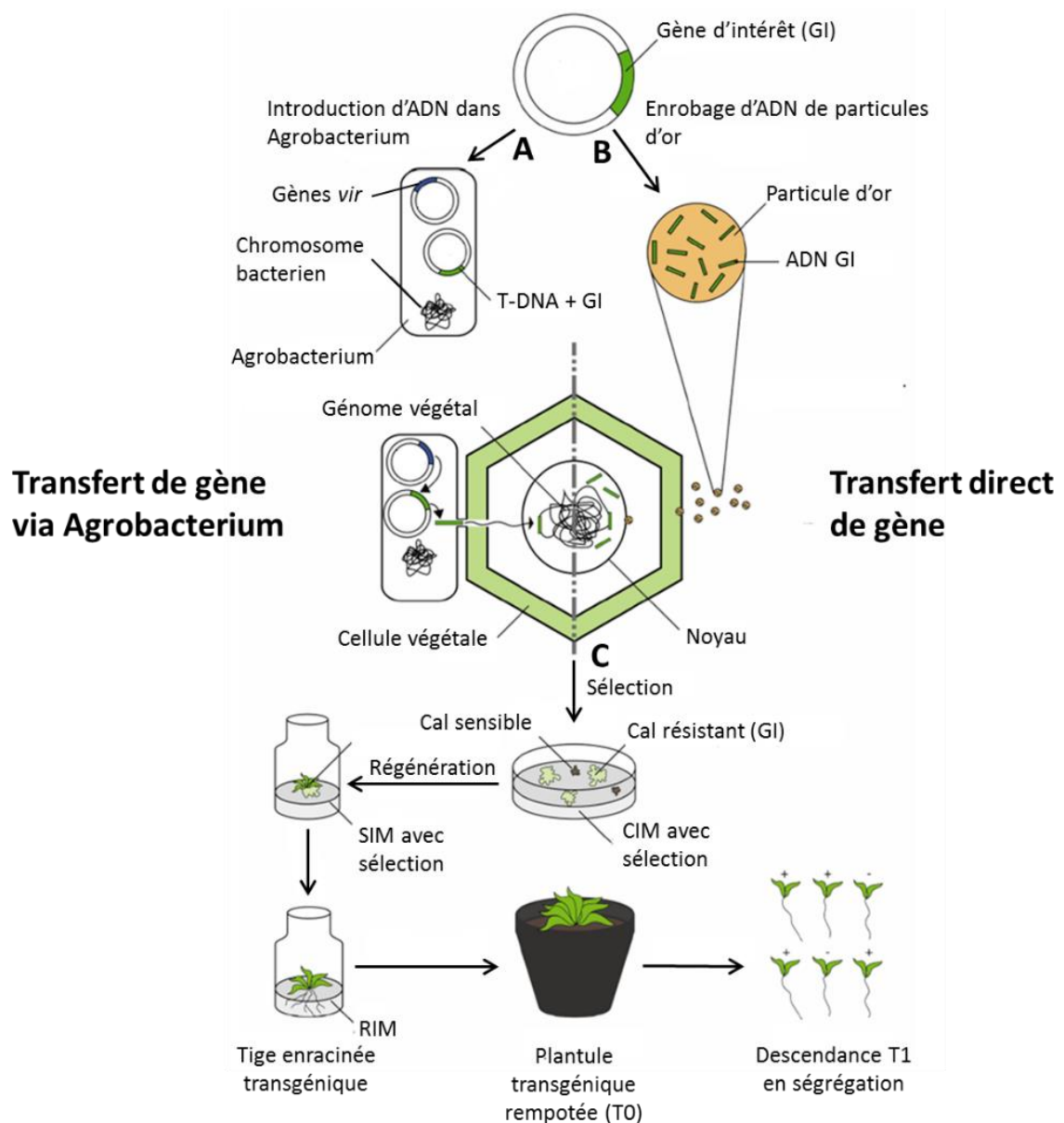


Figure III-2: Techniques de transformation génétique (A) Utilisation d'Agrobacterium comme vecteur d'un transfert T-DNA. (B) Transfert direct d'ADN par bombardement de particules (biolistique). (C) Régénération et sélection des calls et plantules transformés. GI=Gène d'intérêt, Gènes vir=Gènes de virulence Agrobacterium, T-DNA=ADN de Transfert, SIM=Shoot Initiation Medium, RIM=Root Initiation Medium (d'après Anami *et al.*, 2013)

Le transfert d'ADN n'intervient que sur une petite portion des cellules végétales qui subissent la transformation, la sélection des cellules qui ont incorporé le transgène est donc une autre étape clé des techniques de transformation. Ainsi, des gènes marqueurs sont utilisés en parallèle des gènes d'intérêt et sont généralement des gènes de résistance à des antibiotiques comme les néomycines (*nptII*) et hygromicine (*hpt*) phosphotransférases (Fraley *et al.*, 1983; Waldron *et al.*, 1985), des herbicides comme la phosphinothricin N-acetyltransferase (*pat*) et l'acéto lactate synthase (De Block *et al.*, 1989; Wohlleben *et al.*, 1988) ou encore des gènes liés au métabolisme comme la phosphomannose isomérase (Joersbo *et al.*, 1998).

Si *hpt* confère la résistance spécifique à l'hygromycine, *nptII* confère une résistance à une variété d'antibiotiques amino-glycosides comme la kanamycine, la néomycine, la paramomycine et la Généticine. Les gènes de résistance aux herbicides catalysent en une forme inactive le composé actif d'herbicides à spectre large comme le BastaTM. Enfin, la sélection par les enzymes du métabolisme général repose sur un avantage de métabolisation du mannose. Ces gènes sont couplés au(x) gène(s) d'intérêt(s) via des constructions qui assurent une expression constitutive (promoteurs ubiquitaires) afin de pouvoir être exprimés dans une grande variété de types cellulaires (Miki et McHugh, 2004).

1.3.Constitution des constructions et principe de la stratégie CRISPR/Cas9

Les constructions génétiques introduites pour transformer les plantes sont constituées d'un assemblage de gènes comprenant au moins un gène de sélection et un gène d'intérêt, précédés d'un promoteur et suivi d'une séquence terminatrice. De nombreuses variations de cette unité de base existent. Par exemple, une des constructions les plus simples est celle de surexpression qui vise à transférer à la plante une version d'un gène qui sera plus exprimée que la version endogène et exprimée sans contrainte de spatio-temporalité. Cela permet de surproduire une protéine pour détecter son impact sur le phénotype de la plante transformée et d'en étudier le rôle. Ceci est assuré par la séquence régulatrice, le promoteur, issue de différents systèmes comme le gène 35S CaMV (Odell *et al.*, 1985) dont le fort promoteur 35S est utilisé ou les promoteurs de l'ubiquitine (Christensen et Quail, 1996) issus des homologues de différentes espèces d'intérêt comme le maïs (Cornejo *et al.*, 1993; Streatfield *et al.*, 2004) ou switchgrass (Mann *et al.*, 2011). D'autres types de promoteurs peuvent être utilisés pour permettre l'expression spécifique d'un gène sous leur contrôle dans un tissu, organe ou à un stade de développement précis (Dwivedi *et al.*, 2010; Wang *et al.*, 2015). Un

contrôle supérieur de la spatio-temporalité d'expression est permis par l'utilisation de promoteurs inductibles qui sont sensibles à un signal dont la réception entraîne l'activité et l'effet sur l'expression du gène guidé .

Le pendant de la surexpression en transformation génétique est l'invalidation de gènes (KO, Knock-Out) qui interrompt l'expression et/ou la fonction du gène ciblé. Plusieurs méthodes permettent d'aboutir à un KO, parmi elles, l'utilisation d'ADN antisens et siRNA (small interfering RNA) qui interagissent avec les ARNm (ARN messenger) et résultent en une inhibition de la traduction de la protéine correspondante et l'édition directe du génome dont l'intérêt principal est l'édition précise, parfois au nucléotide près, des gènes ciblés. Les techniques d'édition du génome reposent sur l'induction de cassures double brin au locus d'intérêt et sur les mécanismes naturels de réparation de l'ADN, en particulier la réparation non homologue (NHEJ, Non Homologous End Joining) (Jasin et Rothstein, 2013; Manova et Gruszka, 2015) qui est responsable de l'introduction d'erreurs dans la séquence d'ADN réparée. Ce système de réparation imparfait est mis à profit par des techniques de mutations depuis les années 90 avec des enzymes comme des endonucléases, des nucléases en doigts de zinc ou les TALENs (Transcription activator-like effector nuclease) qui induisent les cassures double-brin (Bibikova *et al.*, 2001; Boch *et al.*, 2009; Moscou et Bogdanove, 2009). Une autre méthode d'édition génomique précise est l'utilisation de constructions CRISPR/Cas9 dont l'utilisation a connu une forte croissance après sa mise au point dans la dernière décennie (Chen *et al.*, 2019; Jinek *et al.*, 2012; Wang *et al.*, 2016). Son origine se trouve dans le domaine de la recherche microbiologique puisque la protéine Cas9 et son système de reconnaissance en CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) (Barrangou *et al.*, 2007) sont un mécanisme de défense adaptatif bactérien contre les attaques virales. Grâce à une séquence spécifique de reconnaissance, appelée le « spacer », l'endonucléase Cas9 est capable de cibler un fragment d'ADN contenant la séquence complémentaire à celle du spacer et située en amont d'une séquence appelée « Protospacer Adjacent Motif (PAM) » puis de couper précisément l'ADN (Figure III-3). La révolution du système CRISPR-Cas9 est l'extrême précision et la facilité de mise en œuvre d'un système optimisé par la sélection naturelle. Une autre force du système réside également dans la possibilité d'effacer tout signe résiduel de la transformation. En effet, la modification a lieu à un locus différent du site d'insertion du transgène contenant la Cas9 et les séquences associées, cela permet donc d'éliminer facilement les séquences transgéniques du génome par simple croisement.

Aujourd'hui, la façon la plus simple d'utiliser cette technologie est de cibler un ou plusieurs locus avec l'objectif qu'une mutation délétère y soit induite par la réparation de l'ADN. Ces constructions incluent une (simplex) ou plusieurs (multiplex) séquences spacer de reconnaissance de la ou les cibles sous contrôle d'un promoteur ainsi que la séquence de la protéine Cas9 elle-même sous contrôle d'un autre promoteur. D'autres systèmes reposent sur l'addition d'ADN qui peut s'insérer entièrement au niveau du locus, servir de matrice de recombinaison pour la modification ciblée du locus ou d'autres systèmes qui utilisent la protéine couplée à d'autres domaines activateurs ou répresseurs (Adli, 2018).

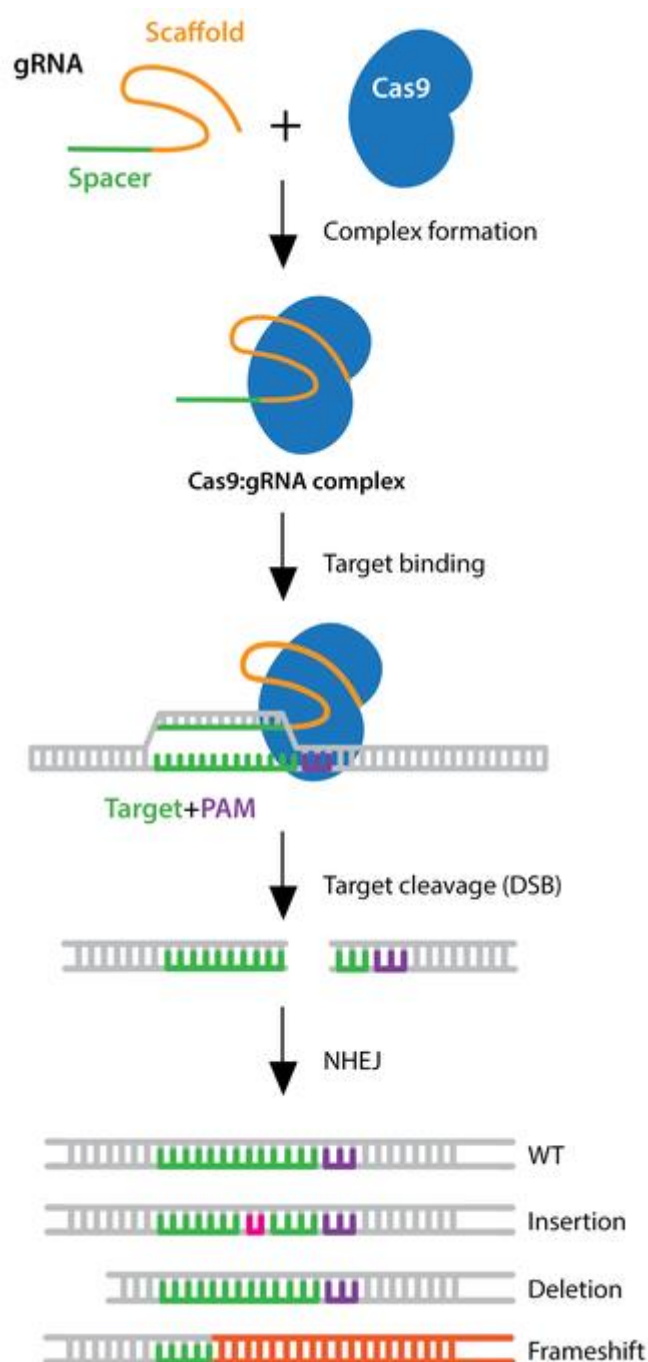


Figure III-3 : Mécanisme de fonctionnement du système CRISPR/Cas9 et de la réparation non homologue (NHEJ). Les gRNA (ARN guide) sont composés d'une structure (scaffold) comportant un « spacer » de reconnaissance de l'ADN cible (target) par la Cas9. Le complexe formé par la Cas9 et l'ARN guide (Cas9:gRNA complex) reconnaît la cible à proximité de la séquence PAM (Protospacer Adjacent Motif) et induit une rupture double brin de l'ADN. Le mécanisme de réparation NHEJ induit une réparation sans modification (WT) ou avec une insertion de nucléotide ou une délétion qui peut modifier le cadre de lecture de l'ADN (Frameshift) (source : <https://www.addgene.org/crispr/guide/>)

1.4. Etat des lieux de la transformation stable chez les graminées et le sorgho

Les premières méthodes de transformation chez les graminées utilisaient dès les années 80 des méthodes de transformation avec des moyens chimiques (polyéthylène glycol, phosphate de calcium) (Krens *et al.*, 1982; Potrykus *et al.*, 1985) et mécaniques (microinjection, bombardement, électroporation) (Fromm *et al.*, 1990; Wan et Lemaux, 1994). Ces méthodes ont permis la transformation de nombreuses monocotylédones dont le sorgho dès 1993 (Casas *et al.*, 1993) avec un taux de succès de 0,28%.

La transformation des dicotylédones est, quant à elle, possible de façon reproductible et efficace depuis 1983 grâce au vecteur bactérien de transfert d'ADN, *Agrobacterium tumefaciens* (Gasser et Fraley, 1989). Les monocotylédones ne répondaient pas aussi bien que les dicotylédones à l'infection par *Agrobacterium* ce qui ne permettait pas d'induire l'expression des gènes de virulence nécessaire à la transformation (Repellin *et al.*, 2001). Après avoir été considérées comme non transformables par cette bactérie, les monocotylédones ont été transformées par *Agrobacterium* 10 ans plus tard que les dicotylédones. Le développement de nouveaux vecteurs chez le riz et le maïs qui expriment un potentiel de virulence plus important dans *Agrobacterium* (Hiei *et al.*, 1994; Ishida *et al.*, 1996) L'addition d'acétosyringone, une molécule activatrice de la virulence d'*Agrobacterium* sécrétée par ses hôtes, à l'inoculation et pendant la co-culture ont également participé à cette avancée (Sant, 2011). Le premier succès de transformation par *Agrobacterium* chez le riz (Chan *et al.*, 1993; Hiei *et al.*, 1994) a été suivi par des succès chez le maïs (Ishida *et al.*, 1996), l'orge (Tingay *et al.*, 1997) ainsi que chez d'autres monocotylédones d'intérêt agronomiques dès les années 90 tandis que la recherche sur le sorgho ne s'est emparée de la transformation par *Agrobacterium* que dans les années 2000 (Gao *et al.*, 2005; Howe *et al.*, 2006; Zhao *et al.*, 2000). Les difficultés rencontrées au début de la transformation génétique chez les graminées expliquent le retard des techniques de transformation chez ces dernières et en particulier pour le sorgho en comparaison aux dicotylédones.

Chez le sorgho, comme chez un grand nombre de monocotylédones, les techniques de culture *in vitro* utilisées reposent principalement sur l'utilisation de l'embryogénèse somatique (Jogeswar *et al.*, 2007; Kaeppler et Pedersen, 1996; Sato *et al.*, 2004). L'induction de cals cellulaires totipotents chez les monocotylédones et leur régénération en plantes entières est un challenge par rapport aux techniques éprouvées chez les dicotylédones. Des protocoles qui

fonctionnent chez certaines espèces de graminées ne fonctionnent pas sur d'autres espèces, même proches. L'unité d'espèce n'est pas non plus suffisante à assurer la transférabilité d'un protocole car l'efficacité peut différer d'une variété à une autre. Le sorgho, en particulier, a la réputation d'être l'espèce monocotylédone de grande culture la plus récalcitrante à la transformation et en particulier à la culture *in vitro* et à la régénération de plantes entières (Altpeter *et al.*, 2016; Gurel *et al.*, 2009; Raghuwanshi et Birch, 2010). Cela pourrait être dû à sa richesse en composés phénoliques qui sont produits en grande quantité en culture *in vitro* et réduisent drastiquement l'efficacité de la transformation (Ahmed *et al.*, 2018; Dykes et Rooney, 2006). Si la transformation d'une variété de tissus (Ahmed *et al.*, 2018) incluant des grains matures a déjà été utilisée (Zhao *et al.*, 2000), les embryons immatures se sont révélés les plus aptes à la transformation (Do *et al.*, 2016a; Gao *et al.*, 2005; Howe *et al.*, 2006; Jeoung *et al.*, 2002; Wu *et al.*, 2014). Les embryons immatures frais sont les explants les plus aptes à la transformation chez le sorgho contrairement à d'autres monocotylédones où des embryons matures secs peuvent être utilisés facilement (Chauhan et Khurana, 2017; Hiei et Komari, 2008). La transformation chez le sorgho est également très sensible au génotype utilisé et la majorité des études utilise les variétés P898012 et Tx430 (Ahmed *et al.*, 2018; Sant, 2011).

Les marqueurs de sélections les plus utilisés chez le sorgho sont *hpt* en transformation par *Agrobacterium* et *bar* en transformation par biolistique. Les autres marqueurs sont *nptII* et *pmi*. Les promoteurs utilisés sont majoritairement ZmUbi (promoteur ubiquitine maïs) et CaMV35S (Ahmed *et al.*, 2018). L'efficacité de ces promoteurs est discutée et ZmUbi semble plus efficace que CaMV35S (Do *et al.*, 2016a; Kumar *et al.*, 2011).

L'efficacité de l'édition des constructions CRISPR/Cas9 est haute chez d'autres monocotylédones comme le riz (jusqu'à 90% en routine)(Ma *et al.*, 2015) mais n'est pas fermement établie chez le sorgho où elle peut s'échelonner de 25 à 80% (Char *et al.*, 2019; Liu *et al.*, 2019a).

Parmi les transformants de sorgho générés et liés à un trait agronomique se trouvent des mutants impliqués dans différents processus biologiques, comme des mutants de résistance aux stress biotiques et abiotiques (Girijashankar *et al.*, 2005; Girijashankar et Swathisree, 2009; Jeoung *et al.*, 2002), de la synthèse et de l'assemblage des kafirines (da Silva *et al.*, 2011; Elkonin *et al.*, 2017; Li *et al.*, 2018a) et des mutants de la biosynthèse des lignines (Scully *et al.*, 2017; Tetreault *et al.*, 2018).

Malgré les succès de transformation par biolistique avec des taux de transformation jusqu'à 25% (Liu *et al.*, 2014; Liu et Godwin, 2012), la transformation par *Agrobacterium* semble plus efficace dans de nombreux travaux (Ahmed *et al.*, 2018). Cette méthode atteint dorénavant des taux de succès de transformation jusqu'à 46% (Belide *et al.*, 2017; Che *et al.*, 2018; Wu *et al.*, 2014) alors qu'ils plafonnaient à moins de 10% il y a une dizaine d'années (Gurel *et al.*, 2009; Kumar *et al.*, 2011).

L'utilisation d'antioxydants, d'agents adsorbants et la culture à l'obscurité ont semblé jouer des rôles importants dans la réduction des effets négatifs causés par les composés phénoliques sécrétés par les cals. En effet, ils peuvent induire la nécrose et la mort des tissus ainsi que des effets toxiques sur *Agrobacterium* (Carvalho *et al.*, 2004; Kaeppler et Pedersen, 1996; Nguyen *et al.*, 2007; Raghuwanshi et Birch, 2010). L'utilisation de plus en plus systématique d'une variété qui produit peu de phénols, Tx430, est probablement également responsable des taux de succès grandissants. L'identification de plusieurs souches d'*Agrobacterium* efficaces pour la transformation du sorgho est également à relever (Jeoung *et al.*, 2002; Wu *et al.*, 2014). Parmi les autres points critiques de l'utilisation du sorgho en transformation mentionnés par Carvalho *et al.* (2004) se trouvent les conditions de croissance et le stade développement des embryons utilisés. En effet, il a été montré que des embryons immatures issus de plantes en champs avaient de moins bonnes capacités de régénération que ceux issus de plantes en serre (Zhao *et al.*, 2000). Ces points sont souvent mentionnés dans les études et leur prise en considération attentive améliore probablement les taux de succès de la transformation.

Malgré les améliorations évoquées ci-dessus, il est toujours difficile de transformer le sorgho comme le démontre le taux de succès de la transformation par *Agrobacterium* qui reste faible dans plusieurs études récentes: 6,2% (Mookkan *et al.*, 2017) et 3,3% (Kuriyama *et al.*, 2019). La maîtrise et l'optimisation du système de transformation du sorgho est donc toujours nécessaire.

1.5.Objectifs généraux des essais de transformation stable

Le premier objectif des essais de transformation mis en place dans le cadre de cette thèse est la maîtrise de la transformation chez le sorgho qui a été testée grâce à deux méthodes. La première est une méthode de biolistique qui obtient des taux de transformation exceptionnellement hauts (Liu et Godwin, 2012). Nous avons eu la chance de bénéficier de l'expertise du groupe de I. Godwin à l'université du Queensland (UQ), auteur de cette publication, grâce à un séjour de 3 mois dans leur laboratoire que j'ai pu effectuer. La deuxième méthode utilisée dans nos essais est plus répandue et repose sur l'utilisation d'*Agrobacterium*. Les différents points relatifs à la maîtrise de la transformation qui seront abordés dans ce chapitre sont l'optimisation des aspects de culture *in vitro*, de livraison d'ADN par les deux systèmes testés et de sélection des plantes transformées.

Le deuxième objectif initial des travaux de transformation stable présentés dans ce chapitre est l'exploration du rôle des gènes candidats de la régulation de la mise en place des parois secondaires décrits dans le chapitre II. Cependant, les difficultés rencontrées n'ont pas permis à cet aspect d'aboutir. De plus, ces travaux ayant démarré avant l'obtention des résultats finaux de l'analyse transcriptomique, nous nous sommes concentrés sur la validation d'orthologues sorgho de FT déjà mis en évidence chez d'autres espèces. Quand la maîtrise du protocole de transformation sera atteinte, il sera possible d'envisager la caractérisation des « nouveaux » FT pour lesquels aucun homologue chez d'autres espèces n'a été mis en évidence pour un rôle dans la régulation des parois secondaires.

1.6.Choix des cibles génétiques pour la transformation stable

Le choix des cibles pour la transformation stable étant intervenu très précocement par rapport aux analyses transcriptomiques et ayant conscience du caractère long et aléatoire de la transformation chez le sorgho, nous avons choisi des gènes cibles dont les rôles dans la mise en place des parois secondaires sont confortés par de nombreuses études chez d'autres espèces. Nous ciblons donc les orthologues chez le sorgho des FT NAC et MYB les plus importants chez *Arabidopsis* (Rao et Dixon, 2018), le maïs et le riz. Ces résultats permettraient d'apporter une preuve de la maîtrise de la transformation du sorgho, qui n'a été menée à bien par aucune équipe en France, ainsi qu'une preuve de la conservation ou de la perte de fonction de ces gènes entre le sorgho et d'autres espèces, tout en donnant accès, par l'étude du transcriptome des mutants, aux acteurs génétiques que ces FT régulent en aval. En

effet, les mutations induites au niveau des acteurs principaux de la régulation ont le potentiel de modifier profondément l'activité des gènes qu'ils régulent.

Plusieurs types de constructions ont été envisagés selon la redondance des relations entre les paralogues ou la facilité à obtenir les constructions. Dans le cas de l'existence de plusieurs paralogues, l'utilisation de plasmides CRISPR/Cas9 entraînant un KO n'est pas envisageable. En effet, les paralogues ont régulièrement des rôles redondants ne permettant pas d'observer les effets phénotypiques de l'interruption de fonction de l'un d'entre eux. Dans ce cas, l'objectif était donc de construire des plasmides de surexpression qui entraînent un phénotype par l'expression augmentée allant au-delà des effets des gènes paralogues et potentiellement redondants. En cas d'un faible nombre de paralogues (deux gènes par exemple) des constructions CRISPR/Cas9 ont tout de même été envisagées en multiplex, c'est-à-dire ciblant à la fois l'un et l'autre des deux gènes. La priorité était cependant donnée aux constructions de surexpression dans les cas où il existait plusieurs paralogues. En l'absence de paralogues, la stratégie était, par contre, de construire des constructions CRISPR/Cas9 simplex afin d'induire des mutations KO. L'identification des homologues chez le sorgho a été réalisée grâce aux analyses de phylogénie décrites en chapitre II.

Plusieurs gènes ont donc été ciblés, correspondant aux homologues des acteurs majeurs de la régulation des parois secondaires chez *Arabidopsis* décrits au chapitre I. Ils peuvent être rassemblés en 6 groupes compte tenu de la similitude de fonction/séquence de certains d'entre eux ; les gènes NST1 et SND1, les gènes MYB42 et MYB85, les gènes MYB46 et MYB83, les gènes MYB58 et MYB63 ainsi que VND7 et VND6 (Figure II-3).

Les quatre premières paires de gènes sont redondantes fonctionnellement chez *Arabidopsis*. Chez le sorgho il existe également deux copies de NST1/SND1 et MYB42/MYB85 mais une seule de MYB46/83. Afin de cibler les homologues chez le sorgho de NST1/SND1 et MYB42/MYB85, deux constructions de surexpression et de CRISPR/Cas9 mutiplex ont été envisagées afin de générer une mutation KO sur les deux gènes en même temps. Pour MYB46/MYB83 une construction en multiplexe a été définie de façon à cibler sur deux sites le même gène cette fois-ci. Ceci augmente les chances de mutation délétère dans la séquence du gène et permet même éventuellement d'espérer une délétion complète de la séquence entre les deux cibles, assurant à coup-sûr la perte de fonction du gène (Zhou *et al.*, 2014a). Pour cibler l'orthologue de MYB46 et MYB83, une construction CRISPR/Cas9 plus simple avec une seule cible, en simplex, a également été définie. Il existe deux orthologues de MYB58 et MYB63 et l'un d'eux est un gène déjà validé fonctionnellement chez le sorgho par

surexpression (Scully *et al.*, 2017). Dans l'objectif d'explorer de nouvelles questions, son paralogue, SbMYB60b, est la cible de nos essais. Afin de s'affranchir de toute redondance entre les gènes SbMYB60 et SbMYB60b nous avons défini une construction de surexpression pour ce gène. Dans le cas de VND6 et VND7 pour lesquels il existe une potentielle grande redondance car ils forment au total un groupe de 7 gènes paralogues chez *Arabidopsis* (Yamaguchi *et al.*, 2011; Zhou *et al.*, 2014b) ainsi que chez le sorgho, trois constructions de surexpression ont été définies (une construction pour l'orthologue de VND6 chez le sorgho et deux constructions pour les deux orthologues de VND7 chez le sorgho). En plus de ces constructions, des constructions CRISPR en simplex et multiplexe de VND6 et VND7 ont été réalisées.

La multiplicité des cibles permet de se prémunir contre l'échec dans l'une des étapes du processus. En effet, à cause de contraintes de temps dans le cadre de la collaboration avec l'équipe de I. Godwin, de difficultés dans certaines étapes des constructions et de difficultés de clonage de certains gènes, probablement à cause de leur faible expression, nous n'avons obtenu que quatre constructions sur les 13 envisagées : une construction surexprimeur de VND7b (oeVND7b), une construction CRISPR/Cas9 multiplexe des orthologues de NST1 et SND1 (mNST), une construction surexprimeur de SbMYB60b (oeSbMYB60b) et une construction CRISPR/Cas9 simplex de l'orthologue unique de MYB46 et MYB83 (sMYB83) (Tableau III-1).

Tableau III-1 : Constructions définies pour l'étude des homologues chez le sorgho (Sb) des gènes majeurs de la régulation de la mise en place des parois secondaires mis en évidence chez *Arabidopsis* (At). Les constructions obtenues sont indiquées en vert. Oe=overexpression/surexpression, s=simplex/construction CRISPR/Cas9 simplex, m=multiplex/construction CRISPR/Cas9 multiplex.

Gènes majeurs chez At	Homologue(s) Sb	Constructions
NST1;SND1	NSTa;NSTb	oeNSTa
		oeNSTb
		mNST
MYB46	MYB83	sMYB83
MYB83		mMYB46
MYB42: MYB85	MYB42a;MYB42b	mMYB42 oeMYB42a oeMYB42b
VND6	VND6	oeVND6 oeVND7a
VND7	VND7a; VND7b	oeVND7b mVND7
MYB58 MYB63	MYB60 MYB60b	oeSbMYB60b

1.7. Etat de l'art la production et la transformation des protoplastes

La transformation stable est un outil puissant pour l'étude du rôle des gènes, cependant, elle n'est pas toujours disponible et optimisée chez toutes les espèces. La difficulté de cette approche rend coûteuses et laborieuses les expérimentations de transformation stable. La disponibilité d'une technique plus simple permettant des résultats rapides est suffisante et donc souhaitable dans de nombreux cas. Avec le développement des technologies « omics » qui donnent accès à de grandes quantités de données à tester, il apparaît nécessaire de pouvoir disposer d'une technique rapide de criblages des gènes à valider. Les systèmes de culture de protoplastes et leur transformation ont ainsi fait l'objet de plus en plus de recherche (maïs (Cao *et al.*, 2014), peuplier (Maćkowska *et al.*, 2014), switchgrass (Burris *et al.*, 2016)). Ils reposent sur l'utilisation de cellules végétales isolées totipotentes débarrassées de leur paroi appelées protoplastes et qui ont la capacité d'intégrer de l'ADN par simple transfection. L'intégration d'ADN par les protoplastes s'apparente à la transfection d'ADN dans des bactéries par choc thermique.

Ce renouveau fait suite au déclin d'intérêt de la culture de protoplastes face aux progrès de la transformation par *Agrobacterium* après avoir été très utilisée au XXème siècle (Cocking, 1985; Davey *et al.*, 2005; Schaeffer et Hotchkiss, 1978; Shimamoto *et al.*, 1989). Les systèmes de culture et de transformation de protoplastes sont apparus particulièrement utiles pour l'analyse de la localisation subcellulaire de protéines (Chen *et al.*, 2015; Nieves-Cordones *et al.*, 2014), des interactions protéiques (Fujikawa *et al.*, 2014; Ma *et al.*, 2015) de réseaux de régulations de l'expression génétique (Nakashima *et al.*, 2014; Pruneda-Paz *et al.*, 2014; Yoo *et al.*, 2007) ou des voies de signalisation (Cao *et al.*, 2014). En particulier, des études ont déjà montré qu'ils peuvent servir à décortiquer les voies de signalisation des plantes et les réseaux de régulation (Asai *et al.*, 2002; Sutter *et al.*, 2005) et qu'ils peuvent permettre d'étudier le rôle de gènes dont l'expression dépend du développement (Kapila *et al.*, 1997).

Dans le cadre de l'étude de la mise en place des parois secondaires, l'utilisation de protoplastes est donc envisageable. Il est cependant important de noter que les protoplastes sont dépourvus de parois, a fortiori de parois secondaires. Seule l'utilisation de construction de surexpression est donc possible. En effet, l'expression forcée d'un gène d'intérêt, même impliqué dans un processus non actif chez les protoplastes, a le potentiel d'induire une chaîne d'expression et de signalisation. L'utilisation de constructions CRISPR/Cas9 qui inactivent un

gène d'intérêt est probablement vouée à l'échec dans le cadre de l'étude des gènes impliqués dans la mise en place des parois, en particulier des parois secondaires.

Les protoplastes ont la capacité de régénérer des plantes entières de la même manière que les explants utilisés pour la transformation stable. Cependant, la régénération de protoplastes de monocotylédones est difficile et a connu un succès limité (Dalton, 1988; Harris *et al.*, 1988; Kyojuka *et al.*, 1987). L'utilisation de la transformation de protoplastes de façon transitoire pour caractériser rapidement des gènes (Burris *et al.*, 2016; Zhang *et al.*, 2011) est l'objectif de cette étude même si la régénération de plantes entières pourra être envisagée dans un deuxième temps. L'analyse haut-débit de gènes est permise par plusieurs autres systèmes biologiques d'expression transitoire comme les feuilles et épidermes de tabacs et d'oignon (Kitajima *et al.*, 2009). Ils sont utilisés en système hétérologue pour tester des gènes d'espèces différentes comme le riz (Lam *et al.*, 2007; Zang *et al.*, 2010), Arabidopsis (Sheen, 2001) et même le sorgho (DeRose *et al.*, 1996). L'avantage des protoplastes par rapport à ces systèmes est qu'il s'agit d'un système homologue puisque les gènes sont testés sur des cellules dérivées de l'espèce végétale cible des analyses (Zhang *et al.*, 2011). Un autre avantage des protoplastes est la possibilité de les isoler à partir de différents tissus végétaux issus de très jeunes plantes, c'est donc un système flexible qui, de surcroît, ne semble pas subir d'effet de la saisonnalité (Mliki *et al.*, 2003).

La transformation des protoplastes a déjà été menée grâce à différentes techniques chez le riz (Page *et al.*, 2019; Uchimiya *et al.*, 1986; Zhang *et al.*, 2011; Zhang et Wu, 1988), le maïs (Fromm *et al.*, 1986; Golovkin *et al.*, 1993) et le sorgho *S. bicolor* et *S. vulgare* (Battraw et Hall, 1991; Hagio *et al.*, 1991; Ou-Lee *et al.*, 1986). La plupart des études remonte à plusieurs décennies. Les méthodes d'intégration de l'ADN exogène dans les protoplastes incluent des méthodes de transformation physiques (biolistique, électroporation, utilisation de Polyéthylène glycol (PEG)) et biologiques (infection par *Agrobacterium*). Les méthodes testées dans les précédents travaux sur le sorgho sont l'électroporation et la biolistique (Battraw et Hall, 1991; Hagio *et al.*, 1991; Ou-Lee *et al.*, 1986). Dans cette étude, nous avons utilisé la méthode de transformation par PEG car son utilisation repose sur un des protocoles les plus simples et une personne expérimentée dans l'utilisation de cette méthode chez le riz était disponible pour formation et appui.

En complément de l'optimisation des protocoles de transformation stables, les objectifs des travaux présentés dans ce chapitre sont d'optimiser la production et la transformation de protoplastes afin de disposer d'une méthode de transformation transitoire facile à mettre en

œuvre. Cet outil sera utilisé pour le test de l'activité des plasmides construits pour des approches plus laborieuses de transformation stable mais aussi pour contribuer à la caractérisation des réseaux de gènes induits.

Dans un premier temps, les conditions de digestion et de purification des protoplastes ont été testées. Elles peuvent contribuer à l'optimisation de la production de protoplastes. Dans un deuxième temps les conditions de transformations et de co-transformation simultanée par deux plasmides ont été testées. En effet, nous envisageons de co-transformer les protoplastes par (i) un plasmide portant le gène d'intérêt et (ii) un plasmide portant un gène rapporteur de fluorescence. Un taux de co-transformation très élevé sera recherché, permettant de s'assurer que la présence du rapporteur renseigne également sur l'intégration du plasmide portant le gène d'intérêt. Enfin, des tests d'extraction d'ARN ont été effectués dans l'objectif de faire des analyses transcriptomiques sur les protoplastes transgéniques.

2. Matériels et méthodes

2.1.Construction des plasmides

2.1.1. Construction des plasmides CRISPR/Cas9

i. Design des spacers

Les séquences codantes (exons hors UTR) de Sobic.008G112200 (SbMYB83, Sobic.007G018100 et Sobic.010G022400 (les 2 orthologues sorgho de NST1) ont été extraites à partir de la base de données Phytozome (<https://phytozome.jgi.doe.gov/>). Les séquences complètes ont été introduites dans le moteur de recherche de spacer crispor.tefor.net ainsi que le PAM de l'enzyme Cas9 qui sera utilisée (NGG). Les séquences proposées par l'outil sont rassemblées dans un tableau généré par le site. Elles sont situées dans la colonne 2 et leur position dans la séquence ciblée est indiquée en colonne 1. Les séquences sont choisies préférentiellement proches du début du gène de façon à éviter qu'une partie soit correctement transcrite et traduite et potentiellement fonctionnelle. Les scores de spécificité de chaque séquence proposée (de 0 à 100, les scores MIT et CFD mesurent le caractère unique de la séquence dans le génome à l'aide deux modèles différents (Doench *et al.*, 2016; Hsu *et al.*, 2013)) est indiqué en colonne 3. Les colonnes les plus importantes pour le choix définitif en dehors de la colonne du score sont la dernière et l'avant dernière colonne qui concernent les off-targets i.e. les cibles non intentionnelles. L'avant dernière colonne décrit le nombre de gènes qui présentent une séquence identique au spacer proposé avec 0, 1,

2, 3 ou 4 bases de différence (mésappariement). Sachant que l'appariement peut se produire même lorsque les séquences sont différentes, ces mésappariements représentent des risques de reconnaissance de gènes non ciblés. Cependant, la position de ceux-ci importe puisque les nucléotides situées juste avant (12 nucléotides) la séquence de reconnaissance Cas9 PAM doivent être strictement identiques au spacer pour être ciblées par la protéine Cas9. Des mésappariements dans cette zone de 12 nucléotides avec la séquence PAM sont donc moins importants que ceux situés au début du spacer. La dernière colonne indique les loci aux séquences problématiquement proches ainsi que la localisation des mésappariements. Les spacers sont donc choisis selon la position dans la séquence codante du gène, son score global, le nombre de off-targets et la position des mésappariements dans le spacer (Bae *et al.*, 2014). La séquence de référence chez le sorgho étant BTX623, une variété différente de celle utilisée en transformation (Tx430). Le génome de Tx430 est séquencé, mais non proposé sur le site de crispor.tefor.net, des vérifications ont été effectuées sur notre plateforme de bioinformatique afin de s'assurer de l'appariement parfait du spacer et de la présence du PAM.

Les séquences d'oligonucléotides spacer déterminées sont commandées de façon à présenter des extrémités cohésives compatibles avec celles générées par la coupure de l'enzyme qui permettra leur insertion dans le vecteur de clonage (*Bsa1* et *Btgz1*).

*GTGT*GCAACTACTACCTCCGCAAGA Sobic.007G018100 *Bsa1* F (mNST)

*AAAC*TCTTGCGGAGGTAGTAGTTGC Sobic.007G018100 *Bsa1* R (mNST)

*TGTT*GGCGTCGGAGCGCATCGACC Sobic.010G022400 *Btgz1* F (mNST)

*AAAC*GGTCGATGCGCTCCGACGCC Sobic.010G022400 *Btgz1* R (mNST)

*GGCA*TACATGCTGCGGAGTGGGCA Sobic.008G112200 *bas1* F (sMYB83)

*AAAC*TGCCCCACTCCGCAGCATGTA Sobic.008G112200 *Bsa1* R (sMYB83)

ii. Introduction dans les vecteurs simplex

Le spacer est intégré par coupure enzymatique *Bsa1* (New England Biolabs ®) (1h 37°C) au vecteur d'entrée riz pU3-sgRNA (Miao *et al.*, 2013, Bes *et al.*, 2019, in press) et sous contrôle du promoteur U3 de la Pol III du riz. Cinq µL de plasmide ont été transformés dans *Escherichia coli* (*E. coli*) de souche DH5α (Subcloning Efficiency™ DH5α Competent Cells, Invitrogen) par choc thermique et ont été sélectionnées sur milieu LB solide avec Kanamycine

(50mg/L). Le contrôle de l'insertion du spacer se fait par digestion enzymatique *BsaI* (1h à 37°C) de l'ADN plasmidique extrait par mini prep (Wizard® Plus Minipreps DNA Purification System, Promega). L'absence de site de restriction indique le succès de la ligation. La séquence du spacer et du promoteur située dans le vecteur d'entrée est intégrée par recombinaison LR (LR Clonase, système Gateway, Thermo Fisher Scientific) au vecteur de destination pOs-Cas9-3.0 (Miao *et al.*, 2013, Bes *et al.*, 2019) contenant la Cas9 optimisée pour le riz sous contrôle du promoteur Ubiquitine de maïs (ZmUbi) et transformé dans *E. coli* DH5α sur sélection Spectinomycine (100mg/L). Le vecteur de destination contient un gène de résistance *nptII* destiné à sélectionner les plantes transformées.

iii. Introduction dans les vecteurs multiplex

La procédure est la même que celle décrite pour le vecteur simplex à la différence que l'insertion des deux spacers se fait dans le vecteur d'entrée pU6-sgRNA qui contient deux sites d'insertion de spacer chacun sous contrôle du promoteur U6 d'une Pol III du riz (Zhou *et al.*, 2014a). L'insertion dans l'un des sites se fait grâce à une coupure enzymatique par *BsaI* (New England Biolabs ®) (1 h à 37°C) et dans l'autre par *BtgZI* (New England Biolabs ®)(1h 60°C). La vérification de l'insertion par *BsaI* se fait par digestion enzymatique par *BsaI* qui doit révéler que le site de restriction n'est plus présent comme pour les vecteurs simplex. La vérification de l'insertion par *BtgZI* se fait, elle, par coupure enzymatique par *BamHI* (New England Biolabs ®)(1 h 37 °C) qui révèle un seul site de restriction si l'insertion est effectuée et deux si le clonage n'a pas fonctionné.

iv. Vérification de l'insertion

La vérification finale du clonage est réalisée avec l'ADN plasmidique extrait (Wizard® Plus Minipreps DNA Purification System, Promega) des *E. coli* transformées et séquencé par l'entreprise Eurofins Genomics.

2.1.2. Construction des plasmides de surexpression pour la transformation par biolistique

i. Clonage des séquences

Une RT-PCR a été effectuée sur 5 µg d'ARN extraits (PureLink® RNA Mini Kit, ThermoFisher Scientific) d'entre-nœuds en développement de sorgho au phénotype sauvage avec l'enzyme SuperscriptTM III RT (Thermo Fisher Scientific) (1 h à 50 °C). Les amorces d'amplification ont été choisies aux extrémités des séquences codantes de Sobic.006G279400

(SbVND7b) et Sobic.006G199800 (SbMYB60b). Les séquences de SbVND7b et SbMYB60b ont été amplifiées par PCR avec la polymérase Kod Hot Start (Merk Millipore) (appariement à 94°C et extension 20 s à 62°C) sur de l'ARN issu d'une RT-PCR d'entre-nœuds en développement de sorgho de phénotype sauvage.

Chacune des amorces est flanquée par une extrémité spécifique aux enzymes de restriction (*NotI* et *PacI*) utilisées pour la ligation dans le plasmide de destination.

GCGGCCGCCATGGATCAGCAGGAGGAGTCT *NotI* F VND7b

TTAATTAAAGGTCAGTCAACAAGGATCC *PacI* R VND7b

GCGGCCGCCATGGGGAAGGGGCGGGCA *NotI* F SbMYB60b

TTAATTAAATCTCACAACCCACATGGTT *PacI* R SbMYB60b

Les séquences des gènes amplifiées aux extrémités franches ont été liguées dans le vecteur Zero Blunt® (Invitrogen) avec la T4 ligase express (ExpressLink™ DNA ligase, Invitrogen) à température ambiante pendant 40 min. Le produit de ligation est transformé dans des bactéries *Escherichia coli* (*E. coli*) de souche DH5α (Subcloning Efficiency™ DH5α Competent Cells, Invitrogen) par choc thermique qui sont ensuite étalées et incubées sur la nuit à 37°C en milieu LB solide avec 50 mg/L de Kanamycine. L'ADN plasmidique est extrait par miniprep (Wizard® Plus Minipreps DNA Purification System, Promega) et la vérification de l'insertion du gène s'effectue par digestion de l'ADN plasmidique avec l'enzyme *EcoRI* (New England Biolabs ®) 1 h à 37°C qui libère l'insert puis migration sur gel d'agarose 2%.

v. Transformation dans le vecteur d'entrée

Après amplification dans *E. coli*, le plasmide Zero Blunt® (Invitrogen) contenant les séquences géniques a été purifié et digéré par 1 h à 37 °C par *NotI* (New England Biolabs ®) et *PacI* (New England Biolabs®) plasmide. Le fragment libéré est purifié (QIAQuick® Gel Extraction Kit, Qiagen) sur gel d'agarose 2 %. Le plasmide de destination pGEM-Ubi-gfp (Liu *et al.*, 2012, Dugdale *et al.*, 1998) a été digéré et purifié de la même manière pour extraire la *gfp* du locus. Le gène d'intérêt est ensuite ligué avec la T4 DNA Ligase (Invitrogen™ ExpressLink™) pendant 40 min à température ambiante. Le produit de ligation est transformé dans des *E. coli* de souche DH5α (Subcloning Efficiency™ DH5α Competent Cells, Invitrogen) par choc thermique qui sont ensuite étalées et incubées sur la nuit à 37 °C en milieu LB solide avec 100 mg/L d'Ampicilline. La vérification de l'insertion du gène s'effectue par digestion de l'ADN plasmidique extrait par mini prep (Wizard® Plus

Minipreps DNA Purification System, Promega) avec les enzymes NotI et PacI (New England Biolabs ®) 1 h à 37 °C et la taille du produit d'insertion est vérifiée par migration sur gel d'agarose 2 %. La vérification finale se fait par séquençage du plasmide avec une amorce F située dans le promoteur ZmUbi par l'entreprise Eurofins Genomics.

2.1.3. Construction et clonage des plasmides de surexpression pour la transformation par *Agrobacterium*

La séquence amplifiée de VND7 a été liguée dans pGEM®-T Easy (Promega) avec la T4 ligase (New England Biolabs®) à température ambiante pendant 2 h et le produit de ligation été transformé dans des bactéries *E.coli* de souche DH5α (Subcloning Efficiency™ DH5α Competent Cells, Invitrogen) par choc thermique qui sont ensuite étalées et incubées sur la nuit à 37°C en milieu LB solide avec 100mg/L d'Ampicilline. L'ADN plasmidique est extrait par miniprep (Wizard® Plus Minipreps DNA Purification System, Promega) et la vérification de l'insertion du gène s'effectue par digestion de l'ADN plasmidique avec l'enzyme *EcoRI* (New England Biolabs ®) 1 h à 37 °C qui libère l'insert puis migration sur gel d'agarose 2%. Le produit de digestion est ligué avec la T4 ligase (New England Biolabs®) dans pCambia2300 (Marker Gene Technologies) digéré avec *EcoRI*. Le produit de ligation été transformé dans des bactéries *E.coli* de souche DH5α (Subcloning Efficiency™ DH5α Competent Cells, Invitrogen) par choc thermique qui sont ensuite étalées et incubées sur la nuit à 37°C en milieu LB solide 50 mg/L de Kanamycine. La vérification finale de l'insertion se fait par séquençage du plasmide avec les amorces M13R et M13F par l'entreprise Eurofins Genomics.

*i. Transformation dans *Agrobacterium tumefaciens* par électroporation*

Un volume de 100 µL de bactéries compétentes *Agrobacterium tumefaciens* de souche EHA105 (Thermo Fischer Scientific) est décongelé sur glace et transformé par électroporation avec 100ng de plasmide. Les bactéries sont resuspendues dans 1 ml de LB et incubées 2 h à 28 °C avec agitation. La culture a ensuite été étalée sur milieu LB solide contenant la rifampicine et l'antibiotique dont la résistance est portée par le plasmide portant le gène d'intérêt et incubée 72 h à 28 °C.

*ii. Transformation dans *Escherischia coli* par choc thermique*

Un volume de 50 µL de bactéries compétentes *E. coli* de souche DH5α (Subcloning Efficiency™ DH5α Competent Cells, Invitrogen) est décongelé sur glace et mélangé avec 1

μL du produit de ligation et incubé sur glace 30min puis placées à 42°C pendant 1min et sur glace 5min. Sont ajoutés 900μL de milieu SOC (Super Optimal broth with Catabolite repression, Thermo Fisher Scientific) puis les bactéries sont placées à incubation 1 h à 37 °C sous agitation (220rpm). La solution de bactéries est étalée sur milieu LB solide avec l'antibiotique de sélection dont la résistance est portée par le plasmide et incubée sur la nuit à 37 °C.

2.1.4. Liste des plasmides utilisés

Pour l'insertion des gènes amplifiés par PCR les plasmides utilisés sont Zero-Blunt™ et pGEM®-T Easy (Figure III-4)(VND7b et SbMYB60b). Le clonage de VND7b a été réalisé dans le plasmide pCambia2300 (Figure III-5) pour l'utilisation en transformation par *Agrobacterium tumefaciens*. Le clonage dans le plasmide pGEM-Ubi-gfp (avec retrait de la *gfp*) a été réalisé pour l'insertion de SbMYB60b et VND7b et l'utilisation pour la transformation par biolistique. Le plasmide pUKN a été utilisé en co-bombardement des essais de biolistique (Figure III-6). Les insertions du spacer de sMYB83 et des spacers de mNST ont été effectuées respectivement dans pOs-sgRNA2.0 et pOs-sgRNA3.0 (Figure III-7A et B). Les ARN guides ont ensuite été introduits dans le vecteur de destination pOs-Ubi-Cas9 par recombinaison pour transformation par *Agrobacterium* et biolistique (Figure III-7C). Les plasmides PGEM-Ubi-gfp (Figure III-6B) et mCherry sous contrôle de CaMV35S sont utilisés pour la transformation de protoplastes et pHK29-p35:GFP (Bes *et al.*, 2019) a été utilisé pour les transformations par biolistique et *Agrobacterium*.

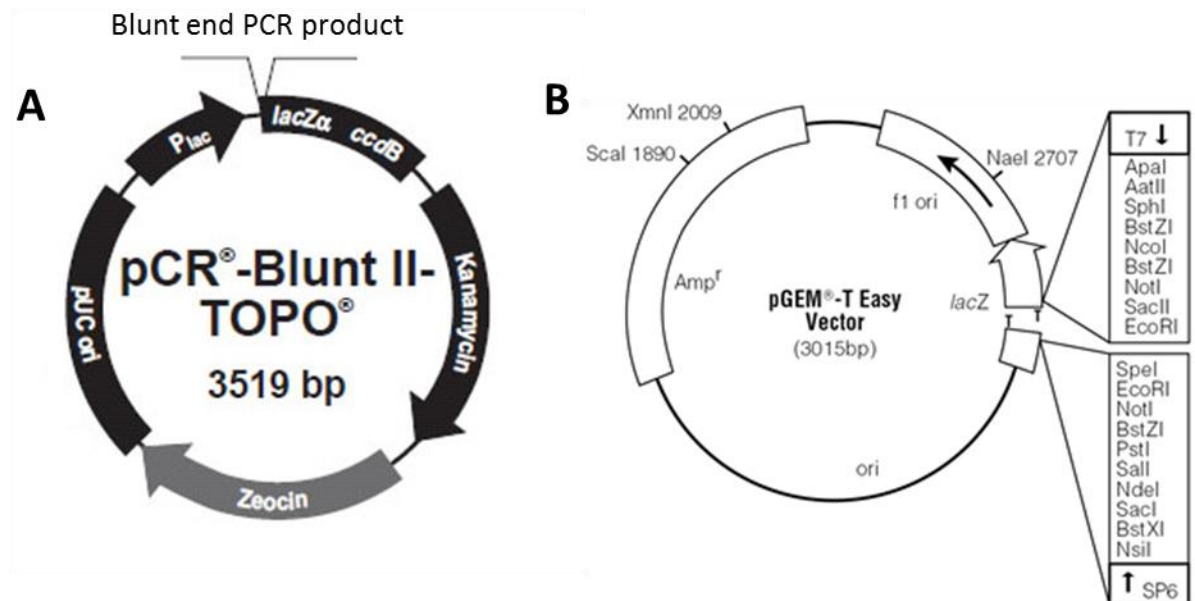


Figure III-4 : Plasmides d'intégration de produits PCR à bouts francs. (A) Zero-Blunt™ (Invitrogen) avec sélection Kanamycine (B) PGEM®-T Easy (Promega) avec sélection Ampicilline. (plasmid.med.harvard.edu et promega.com)

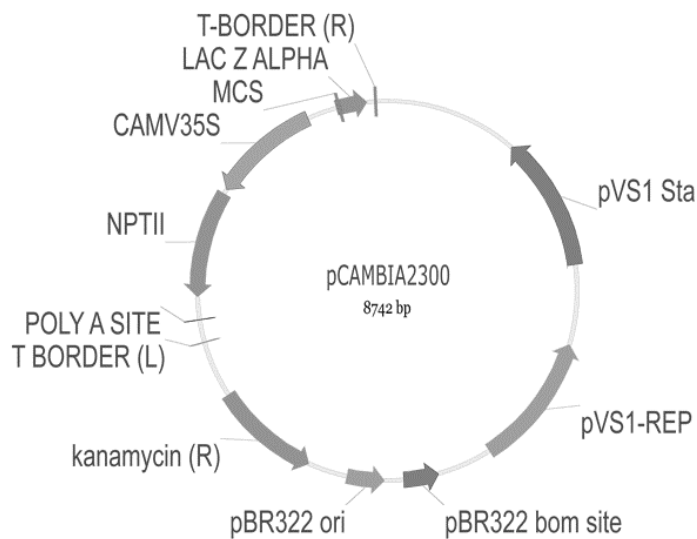


Figure III-5: Plasmide pCambia2300 (Marker Gene Technologies) contenant un site de clonage multiple (Multi Cloning Site, MCS) dans lequel est inséré le gène d'intérêt sous contrôle du promoteur CaMV35S et de la séquence terminatrice CaMV35S poly A. Les extrémités du T-DNA sont représentées par T-BORDER (L) (bord gauche) et T-BORDER (R) (bord droit), la résistance sélection bactérienne est la Kanamycine (*hptII*) tout comme celle de la plante. (markergene.com)

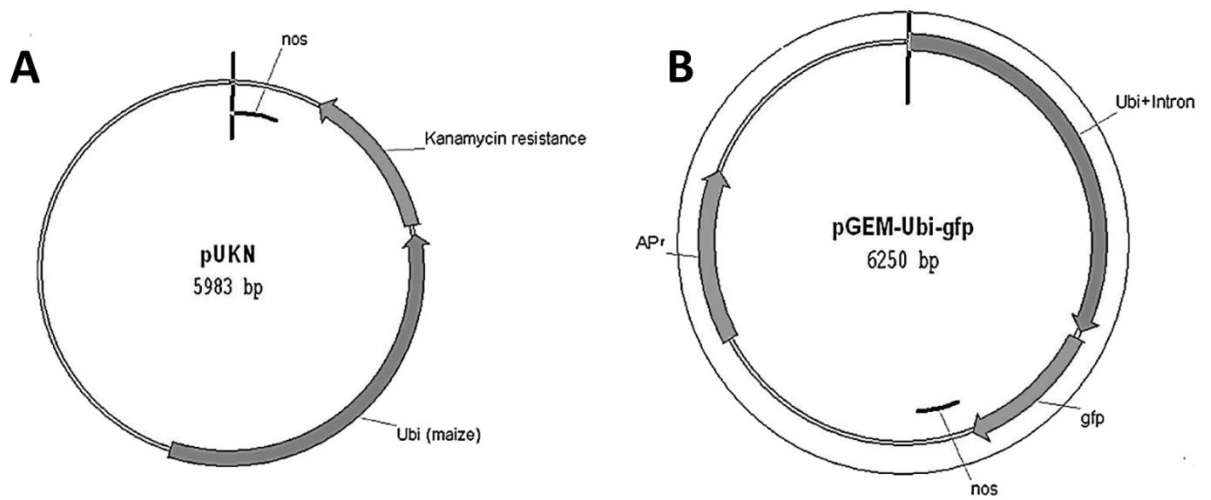


Figure III-6 : Plasmides de surexpression et biolistique fournis par Liu *et al.* (2012) (A) Construction de résistance à la Kanamycine pour co-bombardement avec le plasmide d'intérêt en biolistique (B) Construction de surexpression contenant le gène d'intérêt (ici *gfp*) (Liu *et al.*, 2012)

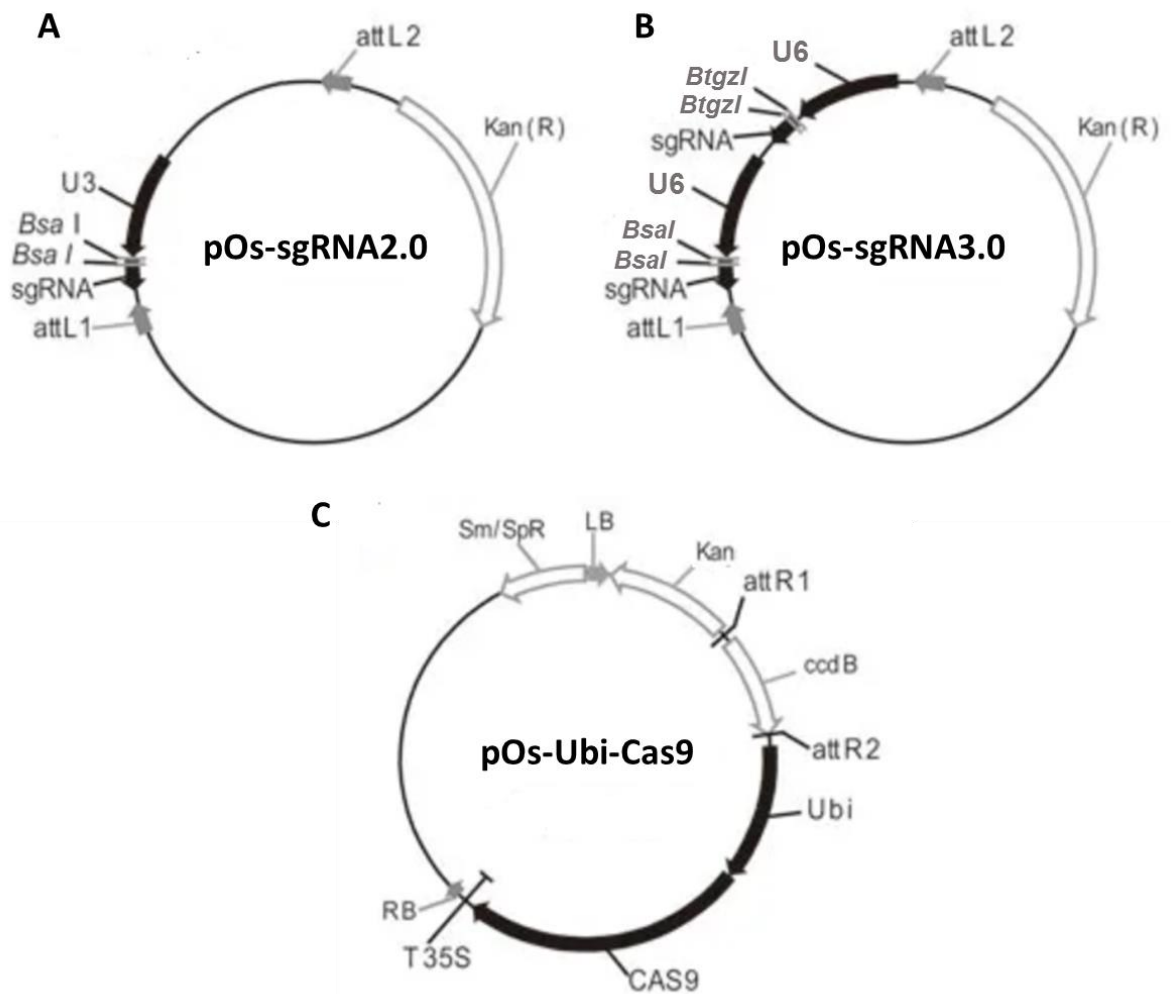


Figure III-7 : Plasmides CRISPR/Cas9 fournis par Bes *et al.* (2019). Vecteurs d'entrée des séquences spacer (sgRNA) simplex (A) et multiplex (B) avec résistance bactérienne Kanamycine (*nptII*) et séquences de recombinaison Gateway indiquées par attL1 et attL2. (C) Vecteur de destination contenant la Cas9 sous contrôle de ZmUbi et terminé par la séquence de fin T35S. Les extrémités du T-DNA sont représentées par LB (Left Border) et RB (Right Border), la résistance de sélection bactérienne est la Spectinomycine (*hptII*) et celle de la plante est la Kanamycine (*nptII*). Les séquences de recombinaison Gateway sont indiquées par attR1 et attR2. Les plasmides d'origine pOs-Ubi-cas9 par Miao *et al.*, 2013 et Zhou *et al.*, 2014 ont été modifiés pour contenir une résistance à la Kanamycine par Bes *et al.* (2019, in press). (Miao *et al.*, 2013)

Nous adressons nos remerciements à Christophe Périn, équipe DAR, unité AGAP, CIRAD Montpellier, qui a fourni tous les vecteurs utilisés pour les constructions CRISPR/Cas9 et à Ian Godwin, Center of Crop Science, QAAFI, Université du Queensland, qui a fourni les vecteurs pGEM-Ubi-gfp et pKUN.

2.2. Transformation stable par biolistique et *Agrobacterium tumefaciens*

2.2.1. Matériel végétal et extraction des embryons

i. Matériel végétal

Des graines de la lignée Tx430 (Miller, 1984) ont été semées en pot et cultivées en serre. Les pots en plastique de 10 L sont remplis d'un mélange en quantités égales de terreau Neuhaus N2 et de terreau Neuhaus S auquel est ajouté de la pouzzolane (1/10 du volume total). Le terreau Neuhaus N2 est un substrat de pH 6 (± 0.3) enrichi en engrais (NPK 14-16-18 1,5kg/m³) constitué d'une tourbe noire de haute qualité qui favorise la ré-humectation et limite le rétrécissement des mottes. Le terreau Neuhaus S contient plusieurs tourbes et son pH varie de 4.0 à 4.5. La croissance des plantes s'est effectuée dans un environnement contrôlé à température de 28 à 30°C le jour et 21°C la nuit avec une hygrométrie non contrôlée. Les panicules ont été recouvertes par des sacs en papiers avant l'apparition des anthères pour éviter des pollinisations croisées.

Les grains sont récoltés entre 12 et 15 jours après la floraison (apparition des anthères). Le stade optimal est déterminé par l'extraction au doigt des embryons qui doivent sortir facilement du grain encore laiteux qui n'a pas encore commencé à former les réserves amidonnées. Une demie-panicule peut être récoltée un jour selon le stade de maturité et le reste les jours suivants.

Les grains de sorgho sont désolidarisés de la panicule et les glumes peuvent être retirées. Ils sont ensuite désinfectés dans de l'éthanol 70% (v/v) pendant 5 min puis dans un mélange d'eau de javel commerciale à 4,8 % (v/v) et de tween 0,2% (v/v) pendant 30min. Ils sont ensuite rincés à l'eau milli-Q stérile 3 fois et enfin trempés 1h dans de l'eau milli-Q stérile avant de subir une deuxième série de désinfection suivant le même protocole.

ii. Extraction des embryons et induction de cals

Les embryons immatures d'une taille entre 0.8 et 1.5 mm sont extraits en condition stérile sous hotte avec une lame de scalpel et isolés sur des boîtes de pétri contenant du milieu CIM (Callus Induction Medium) de Liu *et al.* (2012) (Annexe SIII-1) pour les essais de biolistique et du milieu CIM de Gurel *et al.* (2012) pour les essais de transformation avec *Agrobacterium*. Les embryons immatures sont placés avec le scutellum vers le haut. Les embryons sont cultivés dans une chambre de culture à 27°C à l'obscurité pendant 9 à 11 jours.

2.2.2. Culture in vitro

i. Milieux pour la transformation par biolistique

Tous les milieux de cultures solides (Liu et Godwin, 2012) sont basés sur les milieux Murashike et Skoog (1962)(MS) et contiennent 4.44g/L de poudre MS avec vitamines de Gamborg (Duchefa), 30g/L de saccharose et 8g/L d'agar. Selon le stade de croissance, les milieux (Les CIM, milieux osmotique, milieu de régénération et milieu d'enracinement) ont des compositions variables, décrites en Annexe SIII-1. Le pH des milieux est ajusté à 5.7 avant d'être autoclavé à 121°C pendant 15min et certains composés sont ajoutés de façon stérile après la stérilisation à l'autoclave des milieux. Les milieux de régénération et d'enracinement sélectifs contiennent 30 ou 50mg/L de Généticine (G418).

ii. Culture des cals bombardés

Les cals restent sur milieu osmotique pendant 3 à 4 h à l'obscurité avant le bombardement. Ils sont ensuite transférés sur milieu CIM où ils récupèrent du stress causé par le bombardement pendant 3 jours à l'obscurité à 27°C. Ils sont ensuite repiqués sur milieu de régénération et avec sélection sur lequel ils seront repiqués toutes les deux semaines ou à l'apparition de contaminations et placés en chambre de culture à 27°C et avec une lumière 100µmol/s/m² 16h/jour. Une fois que les plantules ont atteint 1cm elles sont repiquées en milieu d'enracinement avec sélection. Lorsque des racines se sont formées elles peuvent être transférées en serre et plantées dans en pots de 3L.

*iii. Milieux pour transformation par *Agrobacterium**

Les milieux utilisés (Gurel *et al.*, 2012) contiennent tous 4.4 g/L de poudre MS (Duchefa Biochemie, Pays Bas) et sont décrits en Annexe SIII-2. Il s'agit de milieux CIM, de co-culture (liquide et solide), de repos, de régénération sous sélection et d'enracinement. La

carbénicilline à 100 mg/L est utilisée dans les milieux post-co-culture pour freiner la croissance des *Agrobacterium*. La concentration d'antibiotique Généticine (G418) utilisée est de 30 ou 50mg/L.

iv. Culture des cals infectés par Agrobacterium

Après infection par *Agrobacterium* et 3 jours de co-culture en milieu solide, les embryons sont mis en culture à l'obscurité pendant 4 semaines sur milieu d'induction CIM avec sélection 50 mg/L de Généticine (sauf les 10 premiers jours) et 100 mg/L de carbenicilline. Ils sont ensuite repiqués sur milieu de régénération avec sélection en chambre de culture à 27°C avec un cycle de 16 h de lumière par jour et sont repiqués toutes les semaines jusqu'à l'apparition de tiges. Les plantules sont alors repiquées en milieu d'enracinement et transférées en serre pour être plantées dans des pots de 3L une fois les racines développées.

2.2.3. Production d'ADN plasmidique et bombardement des cals

i. Extraction d'ADN plasmidique en grand quantité

Un échantillon des stocks glycérol des bactéries *E. coli* transformées avec les constructions d'intérêt est mis en culture dans 2 ml de LB liquide avec l'antibiotique adéquat (100mg/L d'Ampicilline pour les surexprimeurs et 50mg/L de Spectinomycine pour les constructions CRISPR/Cas9) et incubé sous agitation une nuit à 37°C. La culture liquide est utilisée pour ensemencer 500ml de LB liquide contenant l'antibiotique de sélection qui est incubé sur une nuit à 37°C sous agitation. L'extraction de l'ADN plasmidique est réalisée avec un kit Maxiprep (Plasmid Maxi Kit, Qiagen). L'ADN est resuspendu dans un 50µL à 100µL d'eau à une concentration d'1µg/ml.

v. Préparation des billes d'or

Vingt tubes d'or brut sont préparés simultanément: 50mg d'or sont pesés et suspendus dans 1ml d'éthanol 100% (v/v). La solution est vortexée pendant 5min et laissée à reposer 15 min avant d'être centrifugée avec une centrifugeuse de paillasse à 13 000 rpm pendant 10 s. L'éthanol est éliminé par pipetage et les particules d'or sont lavées trois fois dans de l'eau stérilisée en les vortexant pendant 1 min, laissées reposer pendant 1 min et centrifugées de la même manière que précédemment. Le surnageant est éliminé et les particules sont resuspendues dans du glycérol 50% (v/v). L'or est distribué dans 20 tubes eppendorf à raison de 50µL par tube qui seront conservés à -20°C jusqu'à 1 an. Six bombardements de 6 cals peuvent être réalisés avec chaque tube d'or (0,42 mg d'or par bombardement).

L'enrobage des billes d'or se fait en conditions stériles et sur glace. Le tube d'or est vortexé pendant 5min et 5µg du plasmide d'intérêt et 5µg du plasmide de résistance sont ajoutés à l'or et vortexés 2min puis 20µL de spermidine (0,2M) et 50µL de CaCl₂ (2.5M) sont ajoutés simultanément. La solution est vortexée 2min et précipitée sur glace 5min puis centrifugée à 13000rpm pendant 10s. Le surnageant est éliminé et 130µL d'éthanol 70% (v/v) sont ajoutés puis la solution est vortexée 5min. La solution est précipitée sur glace 5min puis centrifugée à 13000rpm pendant 10s. Le surnageant est évacué et 35µL d'éthanol 70% (v/v) sont ajoutés puis la solution est vortexée 5min. Les particules d'or doivent être utilisées immédiatement après le vortexage.

vi. Bombardement des cals

Des embryons entre 9 et 11 jours d'initiation calleuse ont été utilisés pour le bombardement. Les cals formés qui ont le meilleur potentiel sont choisis par leur taille (3mm), leur forme (bourgeonnante) et leur couleur (blanc opaque). De plus petits cals peuvent être utilisés au risque de ne pas supporter le bombardement. De gros cals transparents ont peu de chance de régénérer des plantes entières. Six à sept des cals choisis sont placés en cercle au centre d'une boîte de pétri contenant du milieu osmotique durant 2 à 3 h maximum à l'obscurité avant le bombardement.

Le bombardement s'effectue grâce à un canon relié à une pompe à vide (Figure III-8) devant la sortie duquel on place des billes d'or recouvertes des plasmides d'intérêt qui sont projetées sur les embryons par projection d'hélium. Le bombardement s'effectue sous hotte après désinfection du canon. Tous les outils sont stérilisés à l'autoclave à 121°C pendant 15min ou par des bains de 30min dans de l'éthanol 70% (v/v).

Des cals non bombardés sont placés sur les mêmes milieux que ceux ayant subi l'essai de transformation ainsi que sur des milieux sans sélection pour être utilisés comme témoins.

Deux machines ont été utilisées, la première, Machine Biolistic PDS 1000, au laboratoire du CIRAD, Montpellier (Figure III-8A) et la deuxième, Particule Inflow Gun (PIG) au laboratoire de l'Université du Queensland, Brisbane (Figure III-8B)

vii. Machine Biolistic PDS 1000.

Un disque de dépôt est placé sur le support prévu à cet effet et 5µL de billes recouvertes de plasmides sont étalés. Un écran d'arrêt et un disque de rupture 1100 psi (indicatif de la pression à laquelle il se rompt) sont installés dans le deuxième support. L'écran d'arrêt sert à

retenir le disque de rupture une fois qu'il a cédé à la pression du gaz projeté. Les supports sont mis en place devant le canon et les cals sont placés dessous dans une boîte de pétri ouverte. Le vide est réalisé jusqu'à 28psi et le gaz est ouvert jusqu'à ce que le disque se rompe, puis le vide est brisé.

La boîte de pétri contenant les embryons est mise de côté à l'obscurité sous la hotte et le processus recommence.

viii. PIG

Le container à gaz est ouvert jusqu'à atteindre une pression de 1000 kPa dans le sas. Un volume de 5 μ L de particules d'or est placé au centre du filtre qui est vissé sur le canon, en haut de la chambre à vide. Une boîte de pétri ouverte contenant les cals est placée sous le canon et la porte est fermée. La pompe à vide est enclenchée jusqu'à ce que la pression monte jusqu'à -90kpa puis la pompe est fermée. La valve de gaz est ouverte pour laisser passer le gaz à travers le filtre et ainsi permettre le bombardement des cals situés en dessous avec les particules d'or.

La boîte de pétri contenant les embryons est ensuite mise de côté à l'obscurité sous la hotte jusqu'à la fin du bombardement de l'ensemble du matériel végétal puis repiqués sur milieu CIM..

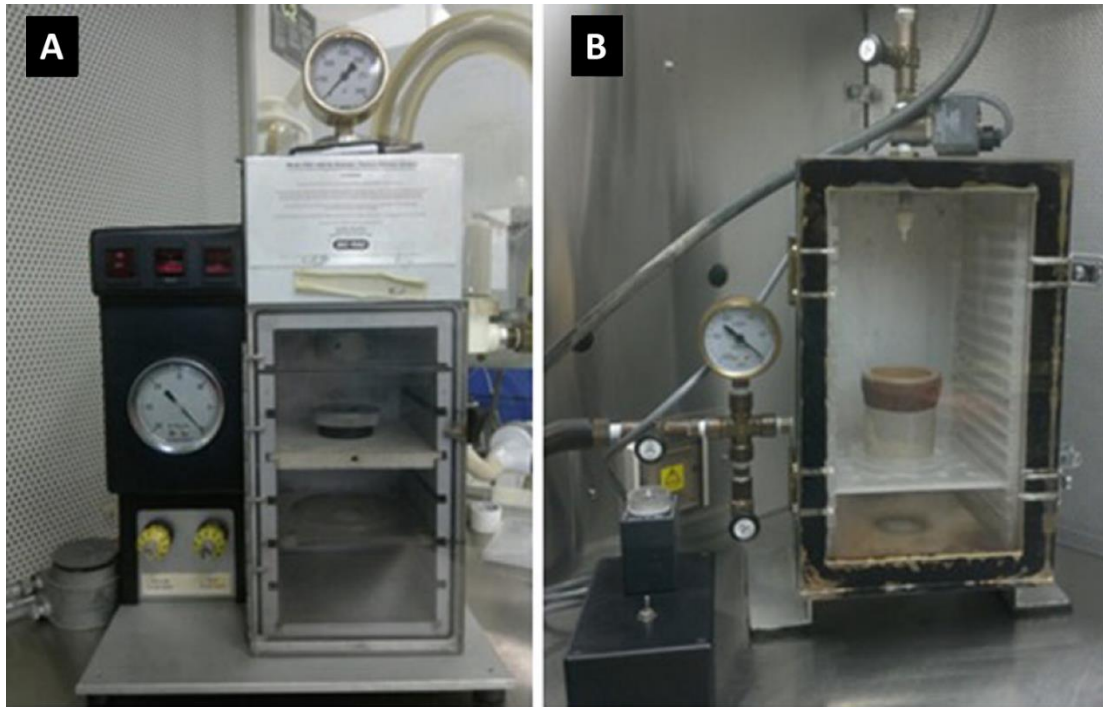


Figure III-8: Canons de bombardement utilisés au (A) CIRAD Biolistic PDs 1000 et à (B) UQ PIG (Particle Inflow Gun)

2.2.4. Infection des explants de sorgho par *Agrobacterium*

Un volume de 10µL du stock à -80°C d'*Agrobacterium* contenant les constructions d'intérêt sont décongelés et étalés sur boîte de pétri contenant du milieu LB solide, l'antibiotique de sélection (50mg/L de Kanamycine ou de Spectinomycine) et 25mg/L de Rifampicine. Elles sont cultivées 2 jours à 28°C puis une colonie est prélevée et mise en culture liquide de 5ml de LB avec antibiotiques sur 1 ou 3 jours. Pour les colonies cultivées 3 jours qui suivent un protocole de virulence exacerbée, une dilution au dixième est réalisée et 10h plus tard les *Agrobacterium* sont centrifugées, et resuspendues dans du milieu AB-MES contenant 200µM d'acétosyringone. La concentration cellulaire est ajustée à une DO de 0,2 avec un spectrophotomètre réglé sur 600nm. Elles sont laissées à agiter pendant 20h à 28°C puis de nouveaux centrifugées et reprises dans 5ml de solution d'inoculation AB-MES/MS avec 200µM d'acétosyringone sans antibiotique ajustée à une DO₆₀₀ de 0,4. Pour les colonies cultivées 1 jour, elles sont suspendues dans du milieu d'inoculation de Gurel *et al.* (2012) et la solution est ajustée à DO₆₀₀=0.4. De l'acétosyringone à 100µM est ajoutée puis les bactéries peuvent être utilisées en co-culture pour l'infection des cals.

Les explants sont mis par groupe de 15 en contact avec la solution d'inoculation ou AB-MES/MS dans un tube eppendorf pendant 5 à 20min avec agitation modérée puis l'excédent d'humidité est aspiré sur papier buvard. Les explants sont placés sur milieu solide de co-culture pendant 3 jours à 24°C à l'obscurité (Gurel *et al.*, 2012).

Des cals infectés avec *Agrobacterium* sans plasmide d'intérêt ou non transformés sont placés sur les mêmes milieux que ceux ayant subi l'essai transformation ainsi que sur des milieux sans sélection pour réaliser une condition témoin.

2.2.5. Génotypage pour validation de la transformation

Le génotypage des plantes peut s'effectuer par détection du gène de résistance *nptII*, de la séquence de la Cas9 en cas de construction CRISPR/Cas9 et en utilisant des amorces sur deux exons différents pour les constructions de surexpression afin de pouvoir différencier la séquence introduite qui ne possède pas d'intron et la séquence endogènes qui en possède. Les amorces utilisées ont été :

GCGATAGAAGGCGATGCG F *nptII*

CCGGCTACCTGCCCCATTCGA R *nptII*

TGCCTGCGGAGGATAGCATGAAGCTC F Cas9

TACCACGAGAAGTACCCGACCATCT R Cas9

TGGATCAGCAGGAGGAGTCT F VND7b

GGGTACTTGCGGTCCTTGTA R VND7b

Des amorces de séquençage ont également été utilisées pour la construction CRISPR/Cas9 mNST :

AGCAAAGCCAGGCAAATA F NSTa (Sobic.007G018100)

CTAGGCAGTTGCATTCACCA R NSTa

CCGATCATTCATTCCCTTTG F NSTb (Sobic.010G022400)

GCATGATCCAGTCGGACTTC R NSTb

2.2.6. Transformation transitoire de protoplastes

Le protocole présenté est adapté du travail de Bes *et al.* (2019) optimisé pour le riz.

i. Préparation du matériel végétal

Un milieu de culture MS/2 (Murashige et Skoog) sans sucres est préparé et ajusté à pH 5.7. Des grains de sorgho matures issus de la variété Texas430 sont désinfectés dans un tube Flacon avec de l'éthanol 70% (v/v) pendant 5min puis dans un mélange d'eau de javel commerciale à 4,8% (v/v) et de tween 0,2% (v/v) pendant 30min. Ils sont ensuite rincés à l'eau milli-Q stérile 3 fois et enfin trempés 1h dans de l'eau milli-Q stérile avant de subir une deuxième série de désinfection. Les grains stérilisés sont mis à germer à l'obscurité à 27°C. Les plantules sont collectées de façon stérile après 7 à 18 jours de culture. Quelques centimètres de tige et de gaine foliaire correspondant à environ 2g sont prélevés dans la partie blanche centrale et réduits en fins morceaux (1 à 2 mm) au scalpel, de façon stérile et à l'obscurité.

ii. Production des protoplastes

L'ensemble des étapes s'effectue en condition stérile, à l'obscurité relative et à température ambiante. Toutes les solutions sont décrites précisément en Annexe SIII-3 et sont ajustées à pH 5.7. Les solutions de PEG, permettant la transformation, et de digestion enzymatique (Macerozyme R10 Onozuka, Duchefa et Cellulase RS Onozuka, Duchefa) sont filtrées et non

autoclavées. Les étapes de manipulations des protoplastes avec pipettes se font avec des embouts coupés en raison de leur fragilité.

Les fragments de tiges sont placés dans une boîte de Petri contenant 20ml de solution mannitol pendant 10min. La solution est filtrée avec une pipette pasteur sur un filtre 70µm. La solution de digestion permettant d'isoler les protoplastes des tissus est versée dans la boîte de Petri contenant les fragments végétaux. Le mélange en digestion est maintenu à l'obscurité pendant 4h (voir Résultats) sous agitation 15rpm et à 27°C. Le mélange est ensuite transféré dans un Erlenmeyer auquel est ajouté un volume égal de solution W5 puis agité pendant 10sec. La solution est filtrée sur une membrane de nylon de 25µm d'épaisseur et les fragments végétaux retenus sont rincés avec 2 fois 15ml de W5. Une deuxième filtration est effectuée dans un nouvel Erlenmeyer. Les solutions contenant les protoplastes sont transférées dans des tubes Corning en verre à fond rond et centrifugées à 100G pendant 10min dans une centrifugeuse à balancier avec accélération et décélération minimales. Le surnageant est éliminé avec une pipette de façon à laisser un large culot. Les différents culots sont regroupés et re-suspendus dans du W5 afin de reconstituer un tube de 13ml. Le tube est centrifugé à 100G pendant 10min dans une centrifugeuse à balancier avec accélération et décélération minimales (2). Les étapes suivantes sont effectuées uniquement en cas d'utilisation de gradient de Percoll qui contribue à la propreté/pureté de la préparation de protoplastes (voir Résultats). Le surnageant est éliminé de façon à conserver 2ml de culot. Des solutions de Percoll à 60%, 40%, 20% et 10% sont préparées en utilisant du W5 comme solvant. Le gradient de Percoll est préparé en ajoutant successivement et doucement le long du tube 2ml de chacune des solutions préparées précédemment en commençant par celle à 60%. Un gradient simplifié avec deux couches de Percoll à 60% et 40% peut être préparé également. Les 2ml de protoplastes sont pipetés et déposés lentement au-dessus du gradient de Percoll sans le perturber. Le tube contenant le gradient de Percoll et les protoplastes est centrifugé à 100G 10min dans une centrifugeuse à balancier avec accélération et décélération minimales (2). Les déchets sont enlevés en pipetant les couches contenant les débris cellulaires (Figure III-12) et les autres couches sont mélangées doucement par inversion. Le tube est divisé en 3 tubes de 13ml qui sont complétés avec du W5 puis centrifugés à 100G 10 min dans une centrifugeuse à balancier avec accélération et décélération minimales (2). Le surnageant est éliminé, les culots sont regroupés et re-suspendus avec du W5 avant de répéter la centrifugation. Les étapes de nettoyage par gradient de Percoll sont terminées à ce stade. Le surnageant est éliminé et le culot est re-suspendu dans 500µL de MMG/MMG2 ou MMG3

(voir Résultats) puis le mélange est laissé à reposer 1h minimum avant la transformation. En cas de transformation le lendemain, le surnageant peut être re-suspendu dans du W5 et le tube laissé sur la nuit couché à 27°C à l'obscurité.

iii. Comptage des protoplastes

Le comptage des protoplastes se fait sur lame de Malassez ou Fush & Rosenthal à partir de la solution de protoplastes/MMG avant la transformation et dans du W5 pour les comptages les jours suivants. Selon la concentration en protoplastes une dilution 1/10 ou plus peut être nécessaire.

Le test de viabilité est réalisé avec du FDA (Fluorescein DiAcétate) (Larkin *et al.*, 1976) qui permet de mettre en évidence par fluorescence les protoplastes viables. En effet, le FDA est métabolisé par les cellules vivantes en un métabolite fluorescent. La solution mère de FDA (5mg/ml) est diluée 100 fois dans de l'acétone. Le FDA dilué est ajouté à la solution de protoplastes à une dilution de 1/20. Après 5min d'incubation à température ambiante 12,5µL sont déposés sur la lame de comptage. L'observation est réalisée sous microscope à fluorescence à la longueur d'onde d'excitation de 440-488nm et d'émission de 520nm. La fluorescence disparaît après 15min.

iv. Transformation des protoplastes

Un volume de 150µL du mélange de protoplastes et de MMG sont placés dans un tube de cytométrie avec 15µg de plasmide à transformer (2 fois 15µg en cas de co-transformation). Le tube est placé 5min à 45°C puis 1min dans de la glace pour appliquer un choc thermique aux protoplastes. 150µL de PEG 40% (v/v) y sont ajoutés doucement avant d'incuber le mélange pendant 15 à 20 min à l'obscurité et à température ambiante. La solution est diluée avec 900µL de W5 et mélangée doucement par inversion. Une centrifugation à 100g dans une centrifugeuse à balancier est réalisée avec accélération et décélération minimales (2) pendant 10min. Le surnageant est ensuite éliminé à la pipette et le culot est re-suspendu dans 50µL de W5. Les tubes sont incubés couchés à 27°C à température ambiante durant la période définie avant comptage ou extraction (18h, 24h, etc.).

v. Extraction d'ARN

Les ARN sont extraits en milieu RNase-free sur une solution fraîche et stérile de protoplastes issue d'une transformation (au minimum 3.10^5 protoplastes). Le kit PureLink™ RNA mini

kit (Thermo Fisher Scientific) est utilisé. Les vérifications de qualité sont effectuées sur un appareil TapeStation en utilisant les kits RNA ScreenTape Analysis (Agilent).

vi. *Tri automatique de cellules (FACS)*

Les protoplastes ont été dénombrés par tri automatique par fluorescence (FACS, Fluorescent Activated Cell Sorting). Ce système repose sur l'utilisation de vibrations afin de briser le flux d'objets à trier en micro gouttelettes séparées. L'utilisation d'un laser permet d'évaluer la taille et la granularité des objets afin de détecter les cellules à considérer. Des mesures de fluorescence peuvent être réalisées sur chacun des objets et ont permis de calculer les taux de transformation et de co-transformation des protoplastes à partir du compte de protoplastes détectés aux longueurs d'ondes d'excitation 490 et d'émission 510 pour la GFP et d'excitation 590nm et d'émission 610 pour la mCherry.

3. Résultats

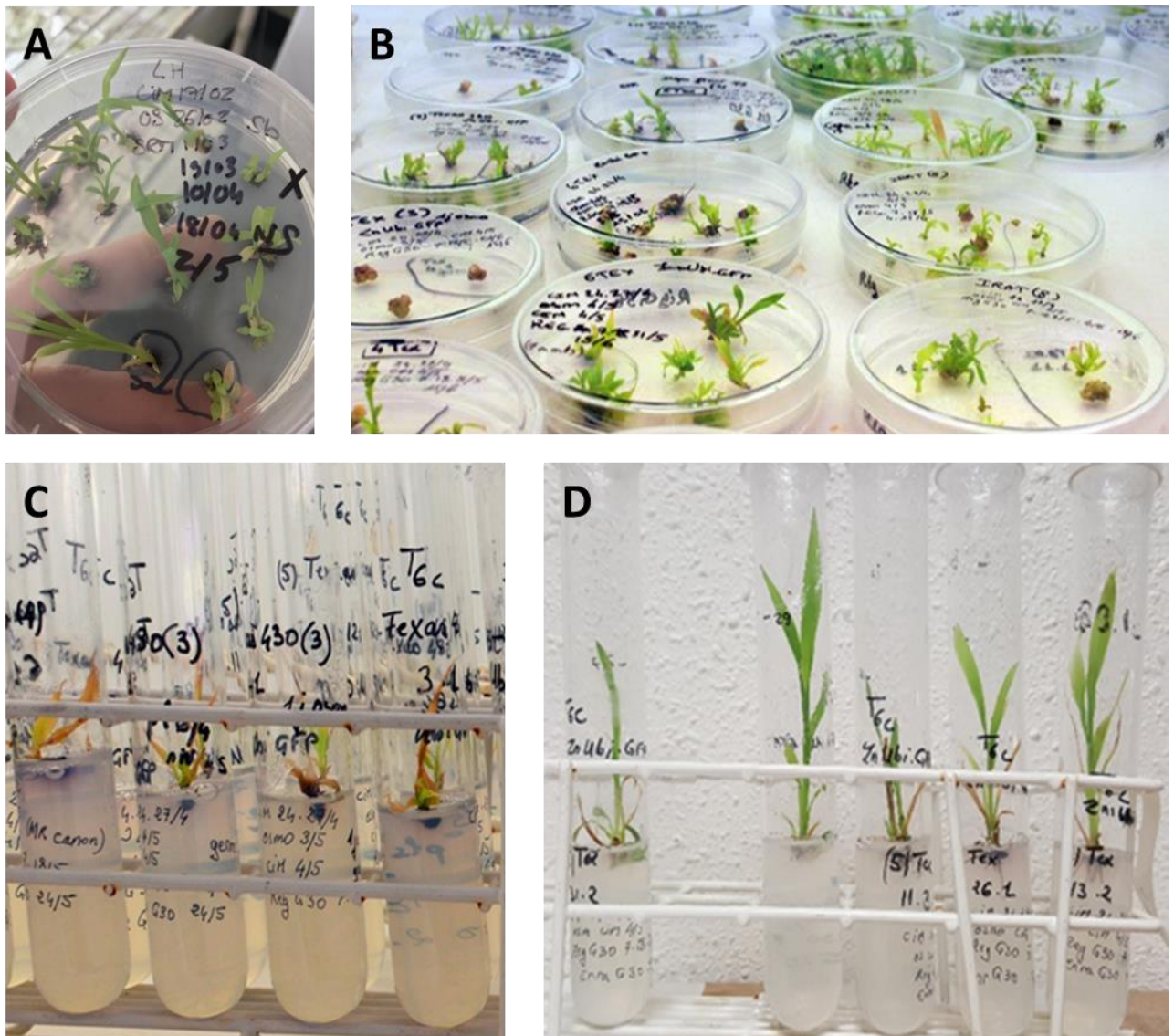
3.1. Essais de transformation stable

3.1.1. Une plante résistante à la Généticine a été obtenue par biolistique

Une première « salve » de transformation selon le protocole de Liu *et al.* (2012) a été réalisée dans les laboratoires de l'université du Queensland (UQ). Cette équipe est pionnière dans la recherche publique du développement de sorghos transgéniques. Elle est la première à avoir largement dépassé les 4% d'efficacité de transformation chez le sorgho en utilisant une technique de bombardement et en obtenant jusqu'à 20% de transformants. Cet essai a été mené en parallèle des essais de K. Massel dont la thèse a été supervisée par I. Godwin.

Deux constructions ont été testées, il s'agit de sMYB83 et oeMYB60b. Ces constructions ont été utilisées en co-bombardement avec un plasmide portant la construction ZmUbi::nptII. Dans un premier temps, 2000 embryons ont été extraits mais seuls 9% ont formé des cals d'embryogénèse. Dix % de contamination par champignon (et levure mais minoritairement) ont pu être observés dans les premiers jours de mise en culture. Une deuxième extraction de 2500 embryons issus de différentes panicules a été menée. Un pourcentage de 90% des embryons ont formé des cals et les plus avancés ont été transformés à J+9 et J+10 après la mise en culture. Environ 750 et 1200 cals ont alors été bombardés avec la construction oeMYB60b et sMYB83 respectivement. Ces chiffres inégaux sont dus à la nature des

constructions. En effet, si la surexpression est assurée dès l'intégration dans le génome à la condition près que le locus d'intégration du transgène soit situé dans une région permettant l'expression du gène, dans le cas du CRISPR/Cas9, le KO est assuré si le transgène est intégré et exprimé et qu'une édition du génome a lieu. Après avoir, pour plus de 90% au cours du premier mois de culture, commencé à émettre une tige, les cals ont été transférés toutes les 2 semaines, ou dès la détection de contamination, dans de nouveaux milieux de génération de tige pendant 3 mois (Figure III-9). Trente-six plantes issues de la transformation avec sMYB83 qui avaient généré des racines en sélection ont été passées en serre au bout de 4 à 5 mois. Aucune plante de la construction oeMYB60b n'a passé le filtre de la sélection. Le génotypage des feuilles de plantes ayant reçu sMYB83 envoyées en France a révélé qu'aucune plante ne contenait la séquence de la Cas9.



Un deuxième essai de transformation par bombardement a été initié en France. Des tests de co-bombardement avec la construction de résistance *ZmUBi::nptII* et une construction fluorescente ont été réalisés sur épiderme d'oignon (*35S::GFP*) et cal de sorgho (*ZmUbi::GFP* et *35S::GFP*) avec le disque de rupture 1100 psi traditionnellement utilisé chez le riz par des utilisateurs de la plateforme génomique du CIRAD et recommandé par Liu *et al.* (2012). Les tests ont été réalisés sur 75 cals par construction fluorescente et deux épidermes d'oignons et les observations ont été faites 24h après le bombardement. Les deux essais avec épidermes d'oignons ont présenté des tâches de fluorescence fortes par rapport à des épidermes d'oignon non transformés. La fluorescence a été observée chez 70% des cals bombardés avec *ZmUbi::GFP* et était bonne en comparaison des épidermes d'oignons bombardés avec *35S::GFP* (Figure III-10). La fluorescence était plus difficilement détectable sur les cals de sorgho bombardés avec *35S::GFP*. Une plante transgénique a été obtenue à l'issue de ces essais, elle contenait seulement le transgène de résistance *ZmUbi::nptII*.

Un nouvel essai de bombardement avec deux transgènes d'intérêt a impliqué l'extraction de 2800 embryons immatures. Mille et 1150 cals ont été co-bombardés avec *ZmUbi::nptII*, et, respectivement, *ZmUbi::VND7B* et *sMYB83*. Soixante-deux plantes *ZmUBI::VND7* et 50 plantes *sMYB83* ont été transférées en serre 6 mois après le bombardement et génotypées. Aucune plante ne contenait le transgène de surexpression ou CRISPR/Cas9 ni le transgène de résistance antibiotique.

Aucune différence de régénération n'a été observée dans la vigueur des plantes entre la chambre de culture à l'université du Queensland et celle de l'UMR AGAP au CIRAD (Figure III-9). De nombreuses feuilles jaunissent au cours de la croissance, ce qui est un signe de stress et de mort très probable. Que ce soit dans le laboratoire de l'université du Queensland ou de l'UMR AGAP au CIRAD, beaucoup de plantules qui avaient continué à se développer normalement pendant plusieurs séries de repiquage ont été sauvées après l'apparition de signes de stress par un repiquage en milieu sans antibiotique afin de relâcher la pression de sélection et leur permettre de croître sans contrainte. Au-delà de ces nombreux cas, quelques dizaines de plantes pour chaque construction testée montraient une coloration verte foncée signe de bonne vigueur même au bout de plusieurs semaines de sélection.

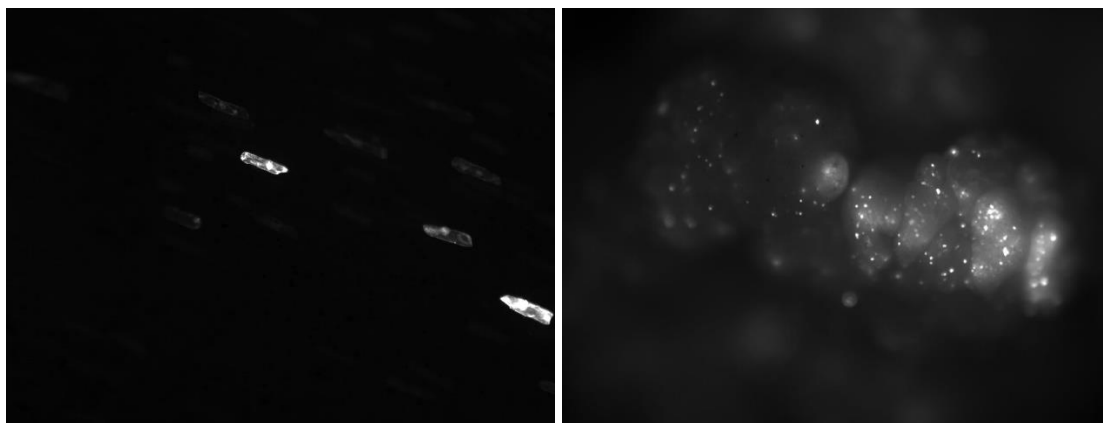


Figure III-10: Fluorescence à 24h après bombardement avec (A) 35S ::*GFP* sur épiderme d'oignon et (B) ZmUbi ::*GFP* sur cal de sorgho de 12 jours

Face aux échecs des essais de biolistic (une plante porteuse du gène de résistance mais sans transgène d'intérêt pour plus de 4000 embryons extraits) et au nombre supérieur de publications récentes utilisant *Agrobacterium tumefaciens*, nous avons commencé des essais de transformation par vecteur biologique.

3.1.2. La transformation par *Agrobacterium* a permis l'obtention d'une plante mutante CRISPR/Cas9 et GFP

Trois transformations ont eu lieu avec vecteur biologique. La première était un test réalisé avec une construction 35S::*GFP*. Un nombre de 319 embryons immatures ont été extraits et mis en culture et 308 cals embryogéniques de 6 jours ont été inoculés par *Agrobacterium*. La fluorescence était détectable les premiers jours après l'infection, mais l'autofluorescence des cals n'a pas permis de la détecter au-delà rendant inutilisable la visualisation de la GFP pour la sélection des tissus transformés. La sélection utilisée de 50mg/L de Généticine a permis de régénérer et enraciner deux plantes dont le génotypage a révélé l'introduction du transgène. Un taux de succès de 0,65% a été atteint lors de cette expérimentation.

La deuxième transformation a impliqué trois constructions (35S::*VND7b*, sMYB83, mNST) pour chacune desquelles environ 400 cals de 6 jours ont été co-cultivés avec *Agrobacterium*. Les plantes régénérées ont été sorties de culture *in vitro* et mises en serre 4 à 7 mois plus tard. Neuf plantes 35S::*VND7b*, 21 sMYB83 et 6 mNST ont été génotypées. Une plante mNST a bien intégré le transgène CRISPR/Cas9. Cependant, le séquençage des gènes cibles n'a pas révélé de mutation. Quarante-huit plantes T1 issues de cette plante transgénique ont été semées en serre et génotypées. Aucune plante T1 ne présentait la séquence de la Cas9, l'édition n'a donc pas été vérifiée. La moitié des plantes T1, voire plus selon l'état de zygote et le nombre de copies intégrées auraient dû être porteuses de la Cas9. Les témoins utilisés lors du génotypage de la plante T0 excluent a priori tout faux-positif mais le locus d'insertion de la Cas9 n'ayant pas été séquencé, cela reste une possibilité. La léthalité de la mutation induite par l'édition de la Cas9 aux étapes de la formation du zygote au développement du grain est également une hypothèse pour expliquer l'absence de détection de la Cas9 en T1. Les gènes NST ciblés par la Cas9 ont cependant été invalidés chez *Arabidopsis* avec un développement apparent normal hormis le défaut de mise en place des parois secondaires dans les tiges (Zhong *et al.* 2011 ou 2008). Des effets de off-target, même si leur faible probabilité a été vérifiée, à la construction du plasmide, peuvent être également en jeu. Globalement, lors

de cet essai, entre 1,5 et 5,3% des cals infectés ont régénéré en une ou plusieurs tiges, cependant un pourcentage $> 0,1\%$ a été effectivement transformé.

La troisième transformation a impliqué les trois mêmes constructions et le même nombre d'embryons extraits que la deuxième. Des cals ont été formés à partir des embryons immatures et co-cultivés avec *Agrobacterium* à 7 jours en moyenne. Un protocole de virulence exacerbée des *Agrobacterium* par addition de $200\mu\text{M}$ d'acétosyringone à leur milieu d'initiation et une sélection de 50mg/L de Généticine ont été appliqués. Des colonies d'*Agrobacterium* se sont développées sur 80% des cals alors que pour les autres expériences, le taux de surinfection était d'environ 15%. Une seule plante a régénéré et a été mise en culture en serre. Le génotypage n'a pas été réalisé car elle n'a pas survécu au passage en serre. Les témoins positifs sans *Agrobacterium* et sans sélection d'antibiotique ont faiblement régénéré, seulement 25% d'entre eux ayant formé des tiges. Il est important de signaler que lors de l'extraction des embryons utilisés pour la dernière transformation les plants de sorgho avaient subi une attaque de pathogènes.

3.2.Optimisation de la production et de la transformation de protoplastes de sorgho

3.2.1. La digestion optimale du matériel végétal est de 4h

Des tests de digestion ont été réalisés afin de déterminer la durée optimale de digestion des parois de cellules des tiges et gaines foliaires chez le sorgho. Le maximum du nombre de protoplastes a été atteint en moyenne avec une digestion de 4h qui a un rendement de 2,6 millions de protoplastes (Mp)/gr de matériel végétal digéré (Figure III-11). Le rendement baisse dès que la durée de la digestion augmente. Cependant l'amélioration du protocole et de l'expertise des manipulateurs a permis d'améliorer le rendement des expérimentations jusqu'à $4,8\text{Mp/gr}$ de matériel végétal (Figure III-12A).

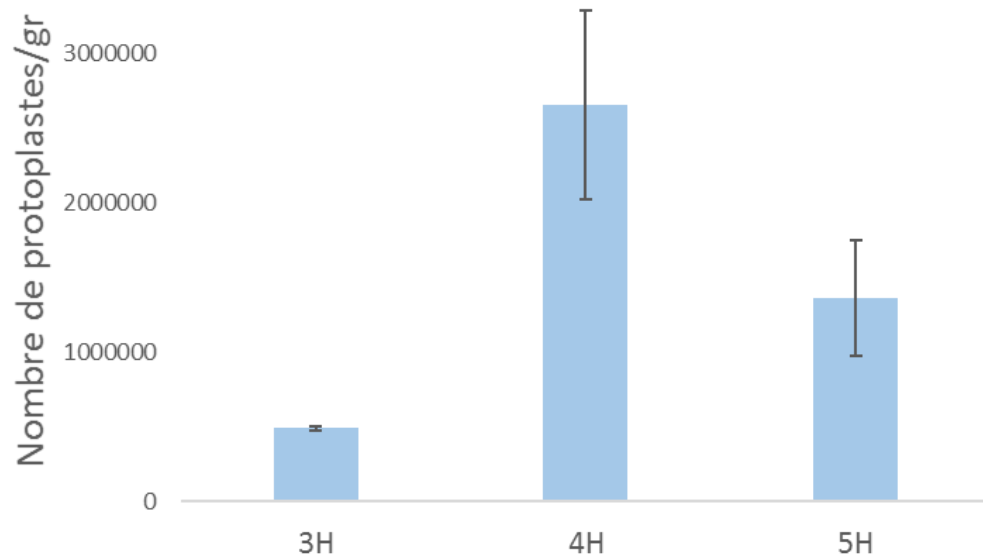


Figure III-11: Nombre de protoplastes obtenus par gramme de matériel végétal en fonction du nombre d'heures de digestion enzymatique (nombre de répétitions : 3h, n=3, 4h-5h, n=4)

3.2.2. La purification de protoplastes par gradient de Percoll n'améliore pas leur qualité

Le gradient de Percoll est une étape délicate où le gradient doit être établi par l'addition minutieuse de couches à différentes concentrations. La centrifugation de la solution de protoplastes dans ce gradient permet de séparer les éléments la composant par densité et d'évacuer les déchets et éléments qui ne sont pas des protoplastes (Figure II-12 B) (Bes *et al.*, 2019). Cette méthode est utilisée en routine au laboratoire pour isoler les protoplastes de riz. Le protocole semble entraîner un plus grand nombre de déchets dans la couche inférieure du gradient de Percoll avec des plantules de riz qu'avec des plantules de sorgho (Figure III-12B et C).

Chez le sorgho, cette étape n'a pas amélioré la qualité des protoplastes produits (Figure III-12A). Le tri et l'utilisation du gradient de Percoll a plutôt tendance à réduire la quantité totale des protoplastes obtenue. Il est difficile de conclure sur l'impact de l'utilisation du gradient de Percoll sur la proportion de protoplastes viables (Figure III-12A).

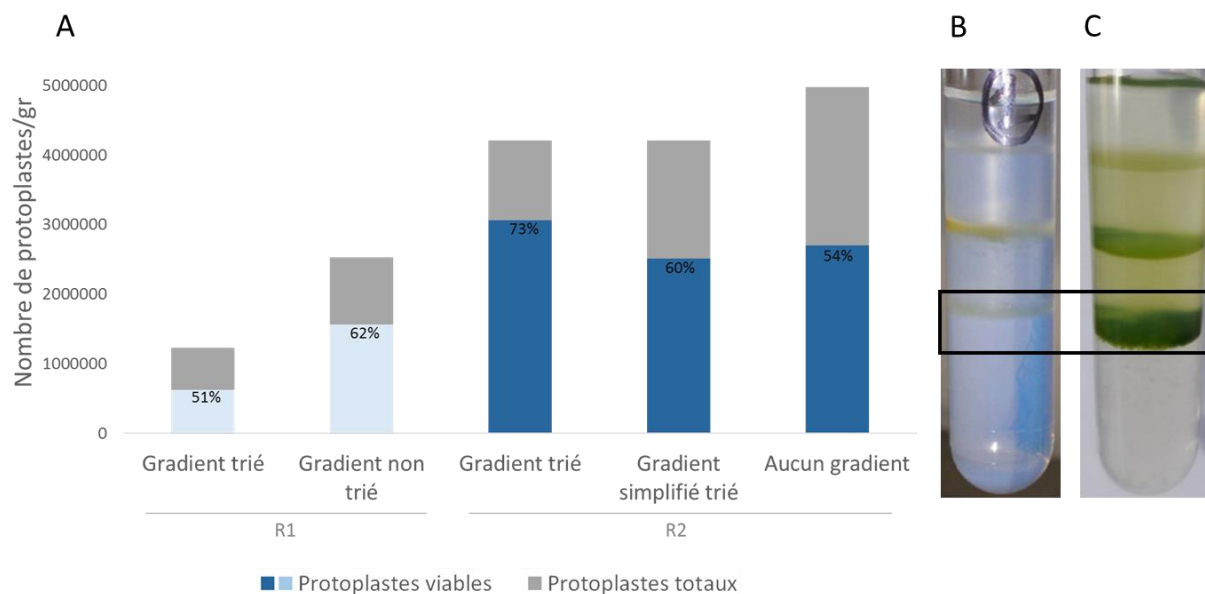


Figure III-12: Impact du gradient de Percoll sur le rendement en protoplastes (A), les portions grises représentent le nombre total des protoplastes et les bleus représentent le nombre de protoplastes viables. Les différentes nuances de bleus désignent 2 répétitions expérimentales indépendantes (R1 et R2). Le gradient de Percoll a été réalisé selon le protocole « riz » pour R1 la première expérimentation de R2 (« gradient trié »). Un gradient simplifié composé de deux couches de Percoll à 60% et 40% a été réalisé dans une condition de R2. Une fois les protoplastes introduits dans le gradient de Percoll, aucun tri des déchets n'a été effectué pour une des deux expérimentations de R1 (« gradient non trié »). Comparaison du gradient de Percoll avec matériel végétal (B) de sorgho et (C) de riz (Bes *et al.*, 2019). La couche de déchets cellulaires est encadrée en noir, les autres couches contiennent les protoplastes à une concentration plus ou moins élevée.

3.2.3. La production de protoplastes de sorgho est variable

Des extractions indépendantes de protoplastes ont été effectuées sur 4 journées d'expérimentation et comportaient 2 répétitions en R1 et R2 (a et b), 3 répétitions en R3 (a, b et c) et 4 répétitions en R4 (a, b, c et d) (Figure III-13). Les nombres de protoplastes totaux et de protoplastes viables ont été mesurés pour chacune des répétitions et leur moyenne par jour a été faite. Les écart-types du total de protoplastes et du pourcentage de protoplastes viables obtenus lors des 4 journées indépendantes ont été calculés et comparés aux écart-types calculés sur l'ensemble de l'expérimentation.

L'écart type du nombre total de protoplastes obtenus entre les jours d'expérimentation présentés ici est d'environ 1Mp/gr ce qui est inférieur à l'écart type du nombre total de protoplastes obtenus au sein même de 3 des journées d'expérimentations sur 4 reportées ici (Tableau III-3). En bref, la variabilité du rendement est plus forte au cœur des différents tubes extraits en parallèle au cours d'une même expérimentation qu'entre des expériences réalisées sur des jours différents.

De même, l'écart type de la viabilité est de 6,4% entre les différents jours d'expérimentation et de 1,8% à 14,1% au sein d'une même journée d'expérimentation (Tableau III-3).

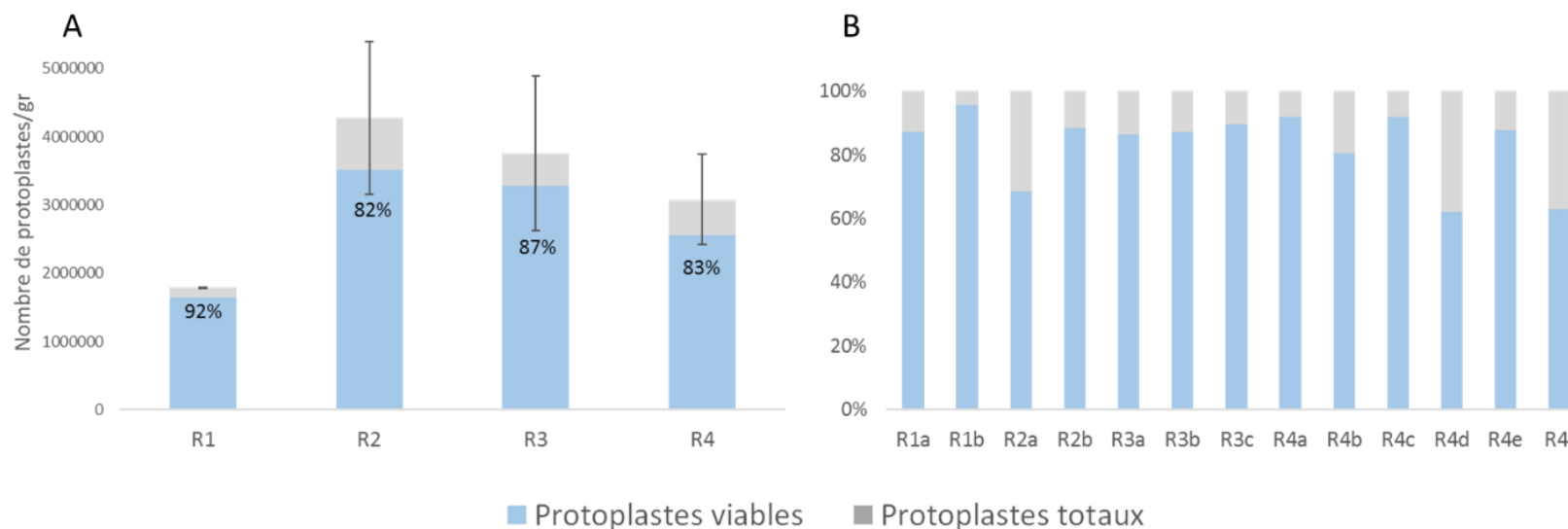


Figure III-13: Variation de la viabilité et de la quantité de protoplastes au cours de 4 répétitions expérimentales. (A) Variation de la viabilité et de la quantité de protoplastes moyenne. Le nombre d'extractions de protoplastes faites pour chacune des répétitions (R1 à R4) est respectivement de n=2, n=2, n=3 et n=6. (B) Viabilité des protoplastes pour chacune des extractions en pourcentage du nombre total de protoplastes.

Tableau III-3: Ecart-types observés entre les répétitions et au sein des répétitions expérimentales pour le total de protoplastes extraits et le pourcentage de viabilité des protoplastes

Ecart-type	Entre jours	IntraR1	IntraR2	IntraR3	IntraR4
Protoplastes totaux	1 076 200	17 678	2 227 386	2 262 788	1 317 518
Protoplastes viables (%)	6,37	6,02	14,10	1,76	13,79

3.2.4. L'aptitude à la transformation des protoplastes est améliorée jusqu'à 75% par l'augmentation de l'osmolarité du milieu et un choc thermique

Le nombre de protoplastes et le taux de viabilité ont été suivis une semaine pour des milieux de reprise dont l'osmolarité était différente. Les taux de transformation et de co-transformation avec GFP et mCherry ont également été évalués à J+2 en conditions d'osmolarités différentes des tampons de reprise et de choc thermique pour la transformation avec un microscope à fluorescence.

La survie des protoplastes semble être améliorée par l'augmentation de l'osmolarité du milieu (Mg^{2+}). En effet, si la quantité de protoplastes extraite est équivalente entre le tampon de base (MMG) et un milieu enrichi en Mg^{2+} (MMG2) (respectivement 1,5 M et 1,7 M de protoplastes/gr), la chute du nombre de protoplastes totaux au cours du temps est plus faible dans le milieu enrichi. Le milieu MMG2 permet la survie des protoplastes durant 7 jours contre 4 jours pour le milieu non optimisé pour le sorgho.

Le taux de viabilité des protoplastes augmente au cours du temps, au contraire du nombre total de protoplastes et de protoplastes viables, pour les deux conditions de milieu et atteint plus de 90% les derniers jours (Figure III-14). Cette augmentation est probablement due à la disparition plus rapide par éclatement des protoplastes non viables parmi les protoplastes totaux (courbe foncée).

Le taux de viabilité atteint 94% dès le deuxième jour pour le milieu enrichi en Mg^{2+} tandis que la hausse de la viabilité est plus lente et la viabilité maximale plus faible dans le milieu de base MMG. Il semble donc que les protoplastes se maintiennent mieux dans le milieu MMG2 et survivent plus longtemps.

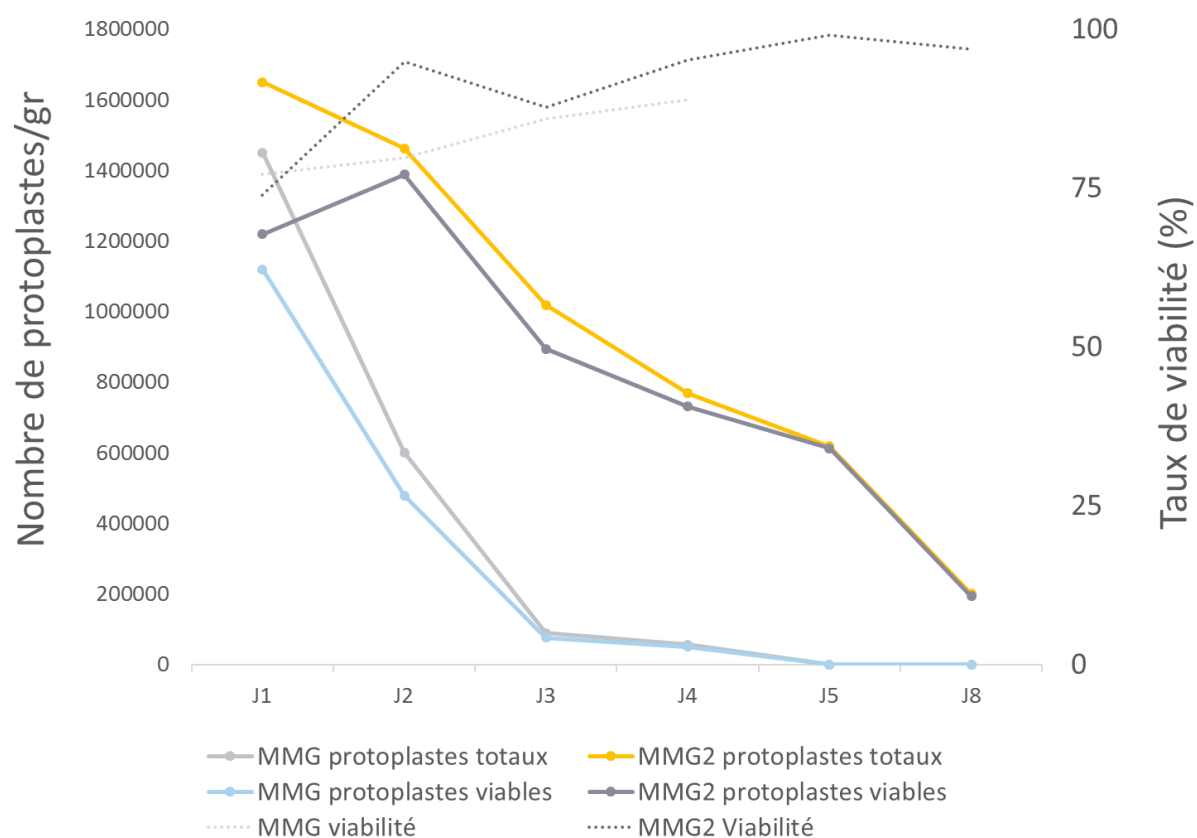


Figure III-14: Nombre de protoplastes totaux, viables et taux de viabilité en fonction de l'osmolarité du milieu (MMG2 est enrichi en Mg^{2+} par rapport à MMG) au cours du temps après extraction

Des tests de co-transformation avec GFP et mCherry ont été effectués avec le milieu MMG2 précédemment testé pour le rendement en protoplastes et avec un milieu à teneur intermédiaire en Mg^{2+} entre MMG ET MMG2 (MMG3)(Figure III-15).

Le taux de transformation semble amélioré par l'utilisation de MMG3 par rapport à MMG2. Ainsi, si des ajustements d'osmolarité sont importants pour la qualité des protoplastes, une trop forte osmolarité semble donc être un frein à leur aptitude à être transformés.

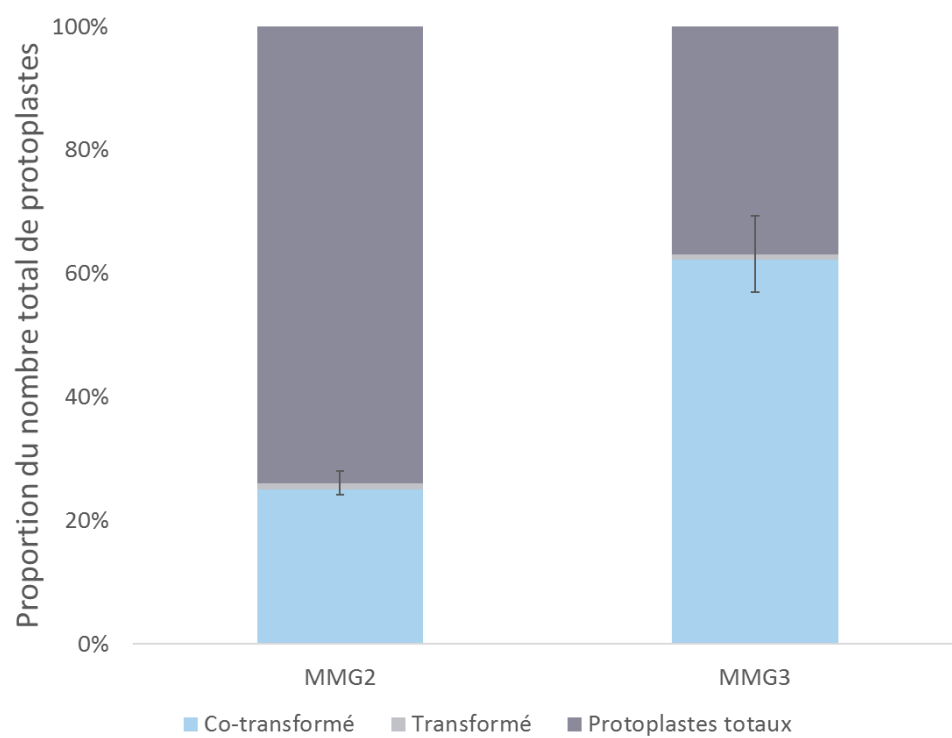


Figure III-15: Taux de transformation et de co-transformation entre les milieux MMG2 et MMG3 (n=4 et n=3). La barre d'erreur indique l'écart-type du % de transformation totale

Un choc thermique a été appliqué juste après la mise en contact des protoplastes et les plasmides GFP et mCherry (Figure III-16). Le taux de transformation total a été évalué pour le protocole avec et sans choc thermique. La répétition R1 implique l'utilisation de MMG2 et les répétitions R2 et R3 celle de MMG3 lors de deux expérimentations indépendantes. Le choc thermique améliore le taux de transformation quelque soit la condition contrôle. Pour les trois conditions testées l'augmentation est de 15% en moyenne. Un maximum de taux de transformation de 75% a été atteint en utilisant un choc thermique combiné à l'utilisation du tampon de reprise MMG3.

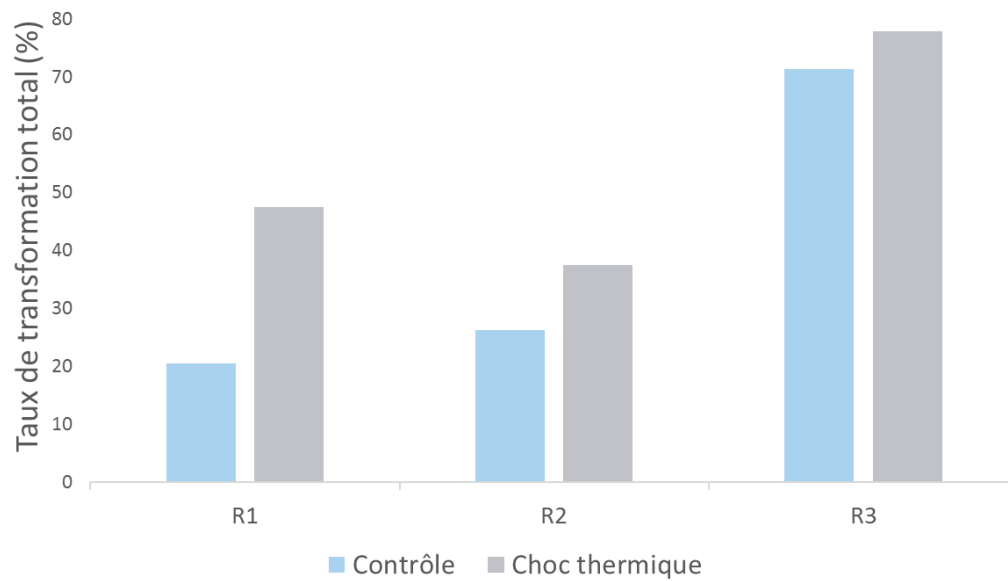


Figure III-16 : Variation du taux de transformation induit par l'utilisation d'un choc thermique. R1 implique l'utilisation de MMG2 et R2 et R3 l'utilisation de MMG3

3.2.5. Le taux de co-transformation de GFP et mCherry est proche 100%

Nous avons transformé les protoplastes avec les protéines fluorescentes GFP et mCherry séparément ou en co-transformation. Après transformation nous avons mesuré au microscope le nombre de protoplastes totaux, le nombre de protoplastes viables puis la transformation sur plusieurs jours. Des mesures au trieur automatique de cellules FACS ont permis d'évaluer le taux de transformation et co-transformation de façon plus précise, avec un comptage indépendant de l'œil humain.

Les mesures de fluorescence après transformation par GFP ou mCherry ont été réalisées au microscope à fluorescence sur plusieurs jours. La fluorescence commence à être émise vers 18h-24h post transformation, que ce soit pour GFP ou mCherry, puis augmente dans un premier temps pour atteindre un pic de fluorescence avant de diminuer. La quantité de protoplastes transformés est maximale pour GFP dès 24h et se maintient jusqu'à 72h post-transformation. La quantité de protoplastes transformés est maximale pour mCherry au bout de 72h ce qui pose la question d'une surestimation du nombre de protoplastes transformés décomptés ce jour-ci. A 24 h post-transformation, la quantité de protoplastes transformés atteint 60% de la quantité maximale théorique de protoplastes transformées avec mCherry (Figure III-17A et B).

La transformation par GFP semble être plus efficace que par mCherry dans un premier temps avec des taux de transformation supérieurs de 9% en moyenne sur les 5 premiers jours. Au cours du temps, le nombre de protoplastes totaux diminue tandis que le taux de viabilité augmente, ce qui explique que le taux de transformation augmente pour les deux constructions GFP et mCherry (Figure III-17C).

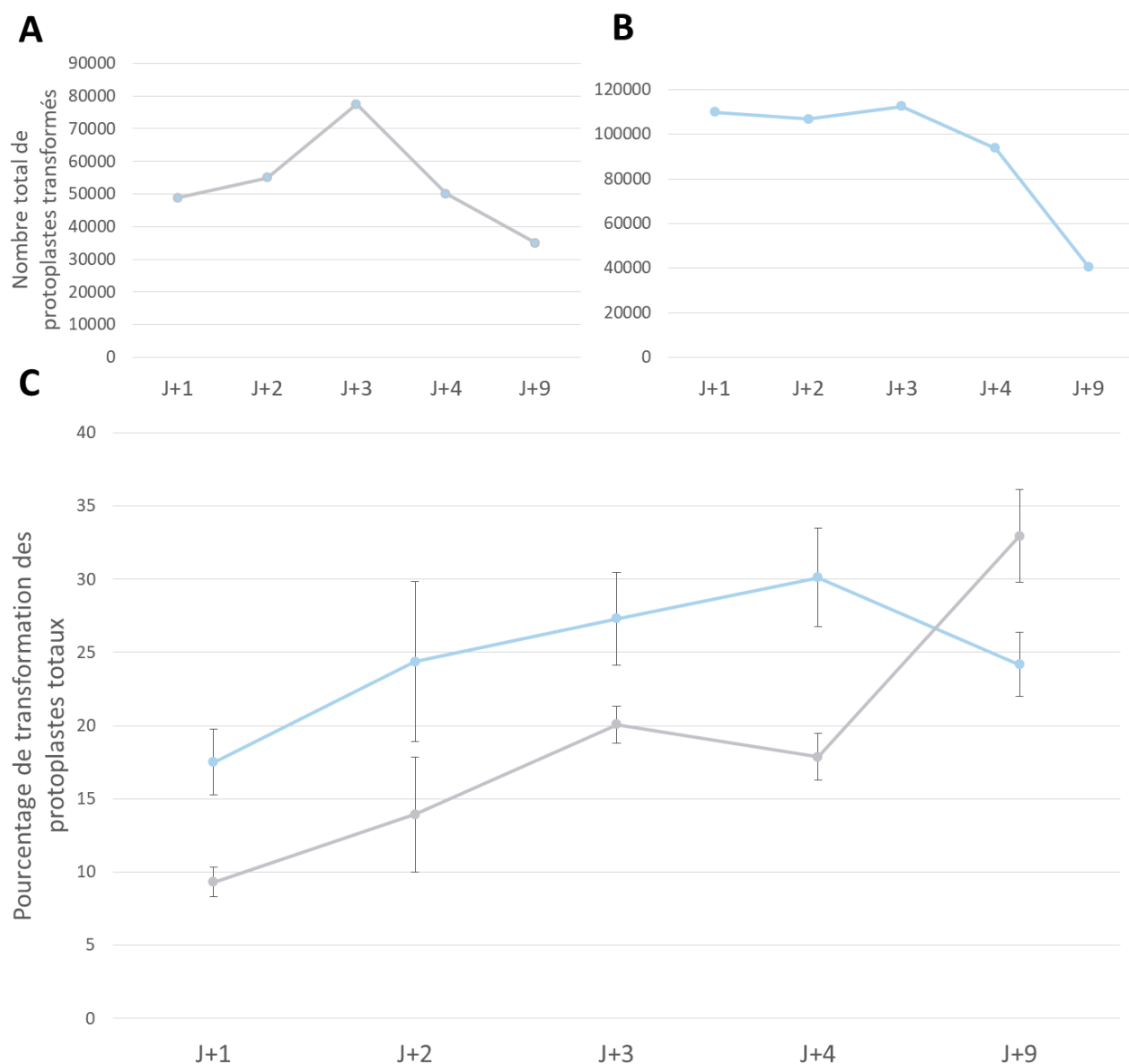


Figure III-17 : Transformation des protoplastes par GFP et mCherry au cours du temps. Quantité totale de protoplastes transformés par mCherry (A) et GFP (B). (C) Taux de transformation par GFP et mCherry (n=4 et n=2) par rapport aux protoplastes totaux

Les observations effectuées au microscope à fluorescence sur des protoplastes co-transformés ont montré que le taux de co-transformation est très élevé (Figure III-18). Sur 20 répétitions, il est en moyenne de 96,3% et au minimum de 91,3%. Il est également très stable ($\sigma=3,7\%$) par rapport aux quantités et viabilités de protoplastes obtenus.

Ces résultats sont confirmés par l'analyse précise au FACS de plusieurs lots de protoplastes co-transformés avec GFP et mCherry. Le comptage de 6008 protoplastes satisfaisant aux critères de qualité (taille, forme, granulosité) (résultat de la filtration de 20 000 objets/événements par le FACS dont 2022 proches de la structure d'un protoplastes mais non sélectionnés) a révélé que 78,5% des protoplastes étaient transformés soit par GFP soit par mCherry soit par les deux et que le taux de co-transformation était de 96% (Tableau III-2).

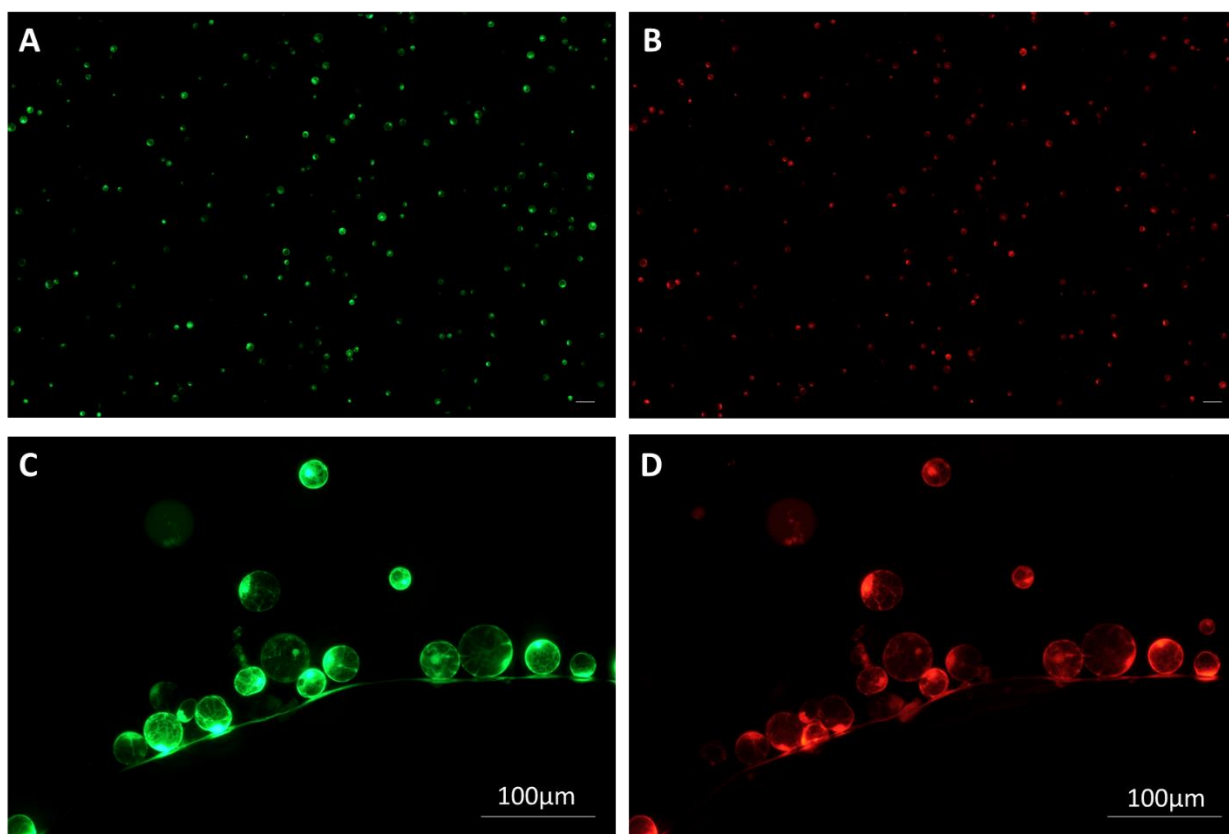


Figure III-18: Protoplastes co-transformés avec GFP et mCherry. (A) et (B) Vue d'ensemble d'une solution de protoplastes transformés et (C) et (D) détails de la fluorescence observée en conséquence de la transformation. (A) et (C) Observation de GFP (Excitation 490, émission 510) (B) et (D) Observation de mCherry (Excitation 590nm, émission 610). Barre d'échelle=100μm.

Tableau III-2: Comptage des différentes populations triées au trieur automatique de cellules à fluorescence (FACS) et taux de transformation et de co-transformation des protoplastes triés

Population		Comptage	Taux de transformation	
Tous évènements		20000		
Protoplastes		8030		
Protoplastes triés		6008		
Protoplastes transformés	mCherry seul	18	78%	0,38 %
	GFP seul	170		3,60 %
	GFP+mCherry	4526		96%

3.2.6. L'extraction d'ARN en quantité compatible avec la création de banques transcriptomiques est possible sur protoplastes de sorgho

Le pic d'expression de la GFP et la mCherry se situe à 72h post transfection mais l'expression est forte dès 24h post transfection. L'extraction d'ARN à 48h est donc optimale pour analyser les effets de la transformation. Les tests d'extraction menés montrent qu'il est possible d'extraire jusqu'à 1,8µg d'ARN par million de protoplastes issus de 2gr de matériel frais à un rendement moyen de 2Mp/gr. Or de 500ng à 1µg d'ARN totaux sont nécessaires pour construire une banque de RNAseq. Les transformations de protoplastes sont donc compatibles en quantité avec la production de données transcriptomiques en aval.

4. Discussion

4.1.Essais de transformation stable

La transformation des graminées a toujours été en retard par rapport à celle des dicotylédones et en particulier chez le sorgho qui est considéré comme une plante récalcitrante à la transformation (Altpeter *et al.*, 2016). Nos essais de transformation par bombardement et *Agrobacterium* ont confirmé l'existence de points critiques lors de la construction de plasmides, du transfert d'ADN et de la régénération.

4.1.1. Choix du matériel végétal

Le choix des panicules et des embryons au bon stade est le point initial de l'expérimentation de transformation pour les deux méthodes. La première expérimentation de biolistique qui s'est déroulée à l'Université du Queensland en présence de manipulateurs expérimentés, a montré que malgré un bon stade apparent, les embryons peuvent être inaptes à la callogénèse (2000 embryons extraits avec moins de 10% de formation de cals). Même si une erreur expérimentale a pu se produire sans être détectée par nos contrôles, ceci illustre la variabilité de réponse à l'induction de cals pour des embryons en apparence au stade optimum.

La variabilité des stades de prélèvement indiqués sur les publications de référence suggère également qu'il est difficile de prédire avec certitude le stade optimal de prélèvement des embryons: 12 à 15 jours après floraison et embryons d'environ 1 mm de long (Liu *et al.*, 2012) pour une transformation par biolistique et 12 à 14 jours après floraison (Gurel *et al.*, 2009), 9 à 12 jours après floraison (Wu *et al.*, 2014; Zhao *et al.*, 2010) pour des

transformations par *Agrobacterium*. Ces incertitudes relèvent probablement du fait que la floraison de l'ensemble de la panicule s'étend sur plusieurs jours mais également de l'état physiologique et sanitaires des plantes. Elles sont un frein à des expérimentations strictement contrôlées.

En complément du stade de développement optimal de l'embryon, il semble qu'un effet de la période de production des embryons existe aussi. En effet, les embryons prélevés en automne et en hiver ont des taux de régénération faibles (jusqu'à un facteur 10 par rapport à ceux prélevés en été) (Do *et al.*, 2016a; Sant, 2011). Les ressources humaines demandées par l'extraction d'embryons (quelques dizaines d'embryons par heure suivant l'expérimentateur) et le suivi des plantes en sélection sont telles qu'il est suggéré de procéder à des expériences de transformation uniquement pendant la période estivale.

4.1.2. Transformation par *Agrobacterium*

Pour la transformation par *Agrobacterium*, les protocoles varient entre co-culture d'*Agrobacterium* avec les embryons immatures directement (Carvalho *et al.*, 2004; Do *et al.*, 2016a; Gao *et al.*, 2005; Gurel *et al.*, 2009; Howe *et al.*, 2006; Kumar *et al.*, 2011; Mookkan *et al.*, 2017; Nguyen *et al.*, 2007; Wu *et al.*, 2014; Zhao *et al.*, 2010) ou avec les cals embryogéniques issus d'embryons immatures (Belide *et al.*, 2017; Elkonin *et al.*, 2017; Jeoung *et al.*, 2002).

L'infection sur cals n'est pas la norme parmi les études sur la transformation par *Agrobacterium* et les travaux qui transforment des cals utilisent des protocoles de bombardement (Belide *et al.*, 2017, Liu *et al.*, 2012) sans longues phases d'obscurité qui s'étendent d'un à plusieurs mois pour les protocoles d'infection par *Agrobacterium* d'embryons immatures frais (Che *et al.*, 2018; Gurel *et al.*, 2009; Kuriyama *et al.*, 2019; Wu *et al.*, 2014). Pourtant nous avons choisi d'utiliser des cals d'une semaine pour nos expérimentations avec *Agrobacterium* et d'y adapter les protocoles de Gurel *et al.* (2009) et Wu *et al.* (2014). En effet, la manipulation d'embryons immatures frais est très délicate et laborieuse. D'une part, il est difficile au cours de la même journée de sélectionner à la serre les panicules au bon stade de développement, de stériliser les graines, d'extraire plusieurs centaines d'embryons et de les mettre en co-culture extemporanément sans mobiliser un nombre important de personnes du laboratoire. D'autre part, la 1^{ère} transformation sur cals d'une semaine avec rapporteur GFP a mené à l'obtention de plantes mutantes avec un taux faible mais relativement comparable à ceux obtenus dans certaines études (Howe *et al.*, 2006;

Kumar *et al.*, 2011; Kuriyama *et al.*, 2019). C'est pourquoi les deux expérimentations suivantes avec les gènes d'intérêt ont été faites sur cals d'une semaine.

Une différence notable entre les deux transformations est la virulence des *Agrobacterium* qui a été augmentée lors de la dernière transformation par 200µM d'acétosyringone contre 100µM pour la première. La plupart des études utilisent 100µM d'acétosyringone (Do *et al.*, 2016a; Nguyen *et al.*, 2007; Wu *et al.*, 2014). Si une concentration plus importante d'acétosyringone est utilisée, elle est répercutée sur la concentration utilisée de carbéniciline ou céfotaxime (Jeoung *et al.*, 2002, Howe *et al.*, 2006). L'infection par *Agrobacterium* peut être responsable de la nécrose de près de la moitié des cals même en condition de succès de transformation par l'endommagement des tissus (Gao *et al.*, 2005; Hansen, 2000; Pu et Goodman, 1992). La surinfection observée sur le dernier essai de transformation par *Agrobacterium* peut expliquer l'échec de la régénération.

4.1.3. Transformation par biolistique

Le 1^{er} bombardement des embryons a été mené avec le matériel et selon le protocole des auteurs du travail de référence (Liu *et al.*, 2012). Si le protocole (composition des milieux, réactifs de préparation des billes d'or, etc.) n'a pas changé pour le 2^{ème} bombardement, l'appareil de biolistique était lui complètement différent. Parmi les différences entre les appareils nous pouvons citer l'utilisation de disques de ruptures avec l'appareil Biolistic PDS 1000 du CIRAD sans équivalent avec celui de UQ et un système de dispersion fabriqué sur par l'équipe de UQ alors que l'appareil du CIRAD possède des grilles standards.

Un contrôle des conditions d'utilisation de l'appareil a été réalisé avec des constructions fluorescentes. La fluorescence observée sur les cals était similaire en terme de distribution sur le cal bombardé à celle observée dans la publication de référence (Liu *et al.*, 2012). Pendant la période de croissance *in vitro*, l'autofluorescence n'a pas permis de détecter de façon ferme la présence d'une fluorescence dans les tissus végétatifs des plantules. Il est probable que la fluorescence observée au cours des premières 24h était le résultat d'une expression transitoire (Gurel *et al.*, 2009). Elle montre cependant que la livraison de l'ADN est effective dans les conditions utilisées.

Les échecs de la transformation par biolistique peuvent venir d'autres aspects variés du protocole. Cette technique semble fortement soumise à l'effet expérimentateur et ce constat de faible reproductibilité n'est pas spécifique au sorgho et a été reporté chez d'autres graminées (Ismagul *et al.*, 2018; Kikkert *et al.*, 2000).

4.1.4. Sélection par Généticine

La sélection par Généticine ne représente qu'une minorité des expériences de transformation du sorgho par *Agrobacterium* (quatre études). Deux études de transformation avec *Agrobacterium* utilisent *nptII* comme gène de sélection et l'une d'entre elles utilise la paromomycine (20mg/L) couplée à la Généticine (10mg/L) (Howe *et al.*, 2016). L'autre utilise une concentration de 250mg de Généticine (Che *et al.*, 2018), ce qui semble difficilement envisageable à cause de l'ordre de grandeur 10 fois supérieur à ce qui est utilisé dans les autres études. Deux études de transformation biolistique utilisent *nptII* en co-bombardement, Liu *et al.* (2012) et Belide *et al.* (2017) utilisent des concentrations de Généticine de 30 et 25-35 mg/L respectivement.

Cependant, l'utilisation de 30 mg/L dans la deuxième expérimentation de biolistique a entraîné beaucoup d'échappées. Ces échappées pourraient également être dues à des repiquages de sauvetage de plantes en milieu sans sélection de façon trop précoce, la sélection pouvant retarder la régénération même en cas de plante de résistance (Cai *et al.*, 1999; Yepes et Aldwinckle, 1994). La sélection non efficace et non franche des plantules de la 2^{ème} expérience de biolistique qui implique un long suivi de la culture *in vitro* nous a conduits à augmenter cette concentration pour les expérimentations suivantes. Au regard des concentrations utilisées dans la littérature, la concentration en Généticine de 50mg/L que nous utilisons est peut-être trop importante pour la régénération de plantules transformées. Néanmoins, les protocoles de régénération étant très différents entre eux, il n'est pas à exclure que la valeur de cette concentration soit moins importante que les moments auxquels elle est appliquée.

4.1.5. Construction des plasmides

La résistance à la sélection des plantules transformées peut être également affectée par une faible expression du gène *nptII*. Pour la transformation par *Agrobacterium* nous avons utilisé le promoteur CaMV 35S pour guider l'expression de *nptII*. Cependant, l'expression des gènes sous influence du promoteur CaMV 35S semble être faible et aléatoire. Le promoteur CaMV 35S est préférentiellement exprimés dans certains tissus (Dong *et al.*, 1991) et les travaux de transformation à partir d'embryons immatures de sorgho de Kumar *et al.* (2011) suggèrent que l'activité du promoteur CaMV 35S est très faible dans les stades précoces du développement des plantules transformées. Do *et al.* (2016) ont également montré que le taux de régénération de plantules sous sélection du gène *bar* avec promoteur CaMV 35 était deux

fois plus faible que ZmUbi (4,1% contre 8,8%) indiquant un potentiel défaut d'expression de la résistance. Cependant, un autre promoteur testé dans cette étude et qui a tout de même permis la sélection de plantes transgéniques, a autorisé plus d'échappées que CaMV 35S, montrant qu'avec un système bien calibré de transformation et de sélection, CaMV 35S peut être utilisé pour sélectionner les plantules transformantes (Do *et al.*, 2016). Ce promoteur a par ailleurs été utilisé dans plusieurs études pour guider différents gènes de résistance de façon satisfaisante (i.e. permettant la sélection de plantes transformées) (Carvalho *et al.*, 2004; Do *et al.*, 2016b; Howe *et al.*, 2006; Kumar *et al.*, 2011; Nguyen *et al.*, 2007).

Pour la transformation par biolistique, les considérations de promoteur ne sont pas justifiées car les constructions sont différentes de la transformation par *Agrobacterium*. En effet, le plasmide co-bombardé avec celui portant les gènes d'intérêt porte le gène *nptII* sous contrôle du promoteur ZmUbi. Celui-ci est actif de façon plus consistante chez les jeunes plantules de sorgho (Do *et al.*, 2016). Cependant, une autre source d'échec de la transformation lié aux plasmides existe pour les essais de biolistique. En effet, si les constructions de surexpression sont de petites tailles et construites selon l'exact modèle de la publication de référence (Liu *et al.*, 2012), les constructions CRISPR/Cas9 bombardées sont incluses dans des plasmides non optimisés pour le bombardement et dans la fourchette haute des recommandations de taille. Elles peuvent donc avoir des difficultés à être insérées dans le noyau et le génome des embryons somatiques.

4.1.6. Conclusion des essais de transformation stable

L'optimisation de la transformation du sorgho étant un long et délicat processus, une conclusion à ce travail ne pourra qu'intervenir au bout de plusieurs années d'expérimentation, et ces études n'ont commencé que depuis deux ans au laboratoire. Cependant, la répétition de ces expériences a permis d'identifier plusieurs points stratégiques d'amélioration.

L'utilisation d'un témoin positif résistant à la Généticine serait un vrai atout pour pouvoir conclure fermement sur les hypothèses qui sont développées ci-dessus. Il permettra de pouvoir optimiser les conditions de culture *in vitro*, et en particulier celles des stades et des doses à employer en sélection lors du protocole de transformation par *Agrobacterium* ou en biolistique pour la sélection. En effet, la culture *in vitro* est une étape cruciale, de nombreuses études portant sur des modifications des phases de culture *in vitro* et de composition des milieux ont été publiées. L'ajout de PVPP au milieu de culture (Zhao *et al.*, 2000) a

notamment permis de bloquer les effets négatifs de la production de composés phénoliques qui est très importante chez le sorgho même si mineure normalement chez la variété Tx430.

Le protocole de transformation étant un processus dont certaines phases sont d'une longueur incompressible, le suivi par étape est nécessaire pour l'optimisation du processus en temps réel sans attendre la régénération des plantes pour leur génotypage. L'utilisation de gènes rapporteurs efficaces est donc importante et il serait idéal d'en utiliser, par exemple un rapporteur GUS ou mCherry, pour pouvoir conclure sur les tests qui seront mis en place.

L'optimisation des conditions de sélection est également un point crucial du succès des expérimentations de transformation. Une trop faible sélection entraîne des problèmes de plantules qui régénèrent de façon trop importante. Le processus de culture *in vitro* demande alors d'intenses efforts de repiquage d'un trop grand nombre de plantes, exacerbés par une très longue phase de sélection quand celle-ci s'éternise. Une expérience de « kill curve » portant sur des écarts de concentration en antibiotique de sélection faibles pourrait donc être mise en place, idéalement avec un témoin résistant à l'antibiotique testé.

Les expérimentations de bombardement seront probablement abandonnées au profit de la transformation par *Agrobacterium*. En effet, les taux de régénération sont dorénavant équivalents à ceux de la transformation par biolistique et reposent sur les travaux de nombreux chercheurs. Le bombardement est de surcroît un processus dont la stabilité de l'intégration et de l'expression du transgène varie. En effet, les loci d'insertion sont difficilement contrôlables au contraire de la transformation par *Agrobacterium*. De plus, les fragments plasmidiques peuvent être intégrés en de très nombreuses copies, portant atteinte à l'intégrité du génome et pouvant mener à des réarrangements de l'ADN et de l'extinction du transgène (Ismagul *et al.*, 2018; Shrawat et Lörz, 2006; Tadesse *et al.*, 2003).

4.2. Production et transformation de protoplastes

4.2.1. La production de protoplastes de sorgho est assurée par un protocole optimisé

Les premières expériences de production de protoplastes avec un protocole optimisé pour le riz ont permis de produire une grande quantité de protoplastes (entre 7.10^5 p/gr et 4 Mp/gr). Ces résultats encourageants ont permis d'envisager l'optimisation d'un protocole de production et transformation de protoplastes de sorgho pouvant être utilisés pour des analyses de cellules transformées de façon transitoires. Si la viabilité des protoplastes était un obstacle

aux expérimentations dans un premier temps, l'optimisation du protocole a permis d'obtenir 1,7Mp/gr dont la majorité est encore viable 96h après leur isolement.

La quantité initiale de matériel végétal suffisante et la durée de digestion enzymatique optimale ont été déterminées. Deux grammes de matériel végétal suffisent à produire une quantité importante de protoplastes. Cela correspond à environ 30 plantules de 10 jours dont les tiges et gaines foliaires non photosynthétiques sont récoltées. Ce sont des quantités utilisées chez le riz (Bes *et al.*, 2019) et du même ordre que plusieurs autres études chez le sorgho (Burris *et al.*, 2016; Maas *et al.*, 1995).

Chez le sorgho, Ou Lee *et al.* (1986) rapportent utiliser 0.8 à 1g de matériel végétal et être en mesure de mobiliser $3-4.10^5$ protoplastes pour des essais enzymatiques. Sairam *et al.* (1999) et Wei et Xu (1993), utilisent, respectivement, des feuilles de sorgho bien développées et des suspensions cellulaires issues de cals et obtiennent des rendements qui varient selon les génotypes testés entre 4.10^3 p/gr et $6,45.10^5$ p/gr. Si les quantités d'enzymes utilisées par Sairam *et al.* (1999) sont similaires à celles que nous utilisons (2% p/v contre 2.25% p/v pour les essais reportés ici), la durée de digestion est de seulement 90 min. La durée de digestion de 3 à 6 h que reportent Ou-Lee *et al.* (1986) est, elle, proche de la durée de digestion optimale de 4h que nous avons déterminée tout en utilisant également des quantités d'enzymes proches (3% p/v contre 2.25% p/v ici). Il semble cependant que l'efficacité du protocole que nous avons optimisé est supérieure à celui utilisé par Ou-Lee *et al.* (1986) et Sairam *et al.* (1999). Même si 2.10^5 protoplastes suffisent à une transformation (Zhang *et al.*, 2011, Burris *et al.*, 2016), l'identification de la durée idéale sur un pas de temps de quelques heures fait partie des améliorations que nous proposons et qui impactent de façon non négligeable le rendement. Par ailleurs, la baisse du rendement avec l'augmentation de la durée de digestion a déjà été observée dans d'autres études (Burris *et al.*, 2016; Chen *et al.*, 2006).

L'utilisation d'un gradient de Percoll pour la purification des protoplastes est un processus délicat et chronophage et qui n'a pas prouvé son intérêt chez le sorgho. Son utilisation réduit potentiellement le nombre total de protoplastes ce qui est probablement dû à la sensibilité de la manipulation des protoplastes avec une pipette. De plus, il n'a pas d'effet évident sur la viabilité des protoplastes.

Cette dernière a pu être améliorée par l'ajustement de l'osmolarité du milieu jusqu'à atteindre 95% au bout de 3 jours de culture. Ces résultats avaient déjà été obtenus par Sairam *et al.* (1999) qui observent seulement 5% de protoplastes morts après 48h. Malheureusement ils

n'utilisent pas les mêmes sels et ne reportent par l'osmolarité de leurs milieux, ce qui ne permet pas de comparer l'origine d'une forte viabilité entre nos travaux.

4.2.2. Obtention d'un taux de transformation de protoplastes supérieur à 70%

L'obtention, la purification, le maintien en vie et la transformation des protoplastes de façon efficace ont donc été permis par des conditions rigoureuses d'asepsie, l'arrêt de l'utilisation d'un gradient de Percoll, une solution de reprise dont la concentration en $MgCl_2$ a été optimisée et l'utilisation d'un choc thermique.

L'augmentation de l'osmolarité de la solution de MMG a permis l'amélioration de la viabilité et de la transformabilité des protoplastes. Cet effet a déjà été décrit dans le cadre d'une transformation avec utilisation de PEG chez le tabac (Negrutiu *et al.*, 1987), le maïs (Armstrong *et al.*, 1990), le palmier à huile (Masani *et al.*, 2014) et switchgrass (Burris *et al.*, 2016). En effet, l'équilibre osmotique est important à la survie des protoplastes dans le milieu et à leur capacité à intégrer de l'ADN exogène.

Ces améliorations ont permis d'obtenir des taux de transformation supérieurs à 70% à plusieurs reprises. Plusieurs études existent sur la production des protoplastes et leur régénération chez le sorgho (Wei *et al.*, 1993, Sairam *et al.*, 1999) ainsi que sur la transformation et la régénération (Battraw et Hall, 1991; Hagio *et al.*, 1991; Ou-Lee *et al.*, 1986). La méthode présentée ici est cependant la première permettant la caractérisation de gènes chez le sorgho en transformation transitoire à l'aide de la co-transformation.

4.2.3. Un taux de co-transformation élevé et stable est obtenu

La stratégie de transformation transitoire que nous avons choisi repose sur la co-transfection d'un plasmide d'intérêt ($ZmUbi::Gène\ d'intérêt$) et d'un plasmide contenant un gène rapporteur codant pour une protéine fluorescente (dans notre cas la mCherry). En effet, la construction d'un plasmide porteur d'un gène d'intérêt et d'un gène marqueur résulterait en un plasmide de grande taille moins facile à intégrer dans les protoplastes. De plus, les gènes d'intérêt que nous souhaitons tester sont déjà inclus dans des constructions de surexpression réalisés dans le même cadre que la construction GFP testée lors de l'optimisation de ce protocole. La co-transformation de mCherry avec les gènes d'intérêt est donc privilégiée.

Un taux de co-transformation supérieur à 95% permet de s'assurer que le taux de transformation détecté grâce au gène marqueur représente fidèlement le taux de transformation du gène d'intérêt. Un taux total de transformation élevé rend possible l'analyse du transcriptome de protoplastes transformés. Ce transcriptome sera comparé à un transcriptome issu de protoplastes témoin, c'est à dire co-transformés avec le vecteur portant la mCherry et celui portant la construction ZmUbi ::*GFP*.

Le taux de co-transformation atteint en routine grâce à l'optimisation de ce protocole est supérieur à 95% et peut atteindre 99%. Ce résultat est une condition absolument nécessaire pour les caractérisations de la fonction des gènes d'intérêt via cette approche et nous permet donc de l'envisager.

Des quantités d'ARN importantes ont également été extraites des protoplastes. L'utilisation des protoplastes dans le cadre de l'analyse transcriptomique de gènes est donc envisageable.

4.2.4. La production de protoplastes reste soumise aux variations de l'environnement

Si le taux de co-transformation est très stable, la quantité de protoplastes, leur viabilité et le taux de transformation restent variables. En effet, les variations ne se limitent pas à l'expertise du manipulateur, même si celle-ci est primordiale pour l'obtention de bons rendements. La sensibilité des protoplastes aux conditions environnementales fait varier drastiquement le succès de leur extraction et leur transformation. A chaque transformation le taux de transformation par le gène rapporteur est vérifié afin de s'assurer d'un taux de transformation compatible avec des études complémentaires (études d'expression par exemple).

Plusieurs points de sensibilité du protocole ont été mis en évidence. Les cônes de pipettes utilisés pour transférer les solutions contenant les protoplastes doivent être coupés. La centrifugation doit être la plus douce possible avec des phases d'accélération et de décélération progressives. Les manipulations doivent être minutieuses mais aussi rapides pour limiter la durée totale de l'expérimentation. L'asepsie est importante dans le cadre de suivi d'expression car une contamination microbienne réduit la quantité et la qualité des protoplastes et perturbera l'analyse du transcriptome. Certains lots de graines, pourtant désinfectés, peuvent être contaminés et disposer de plusieurs lots de graines indépendants est important.

Discussion générale

1. Rappel des objectifs et des stratégies

Les parois secondaires contribuent à de nombreux rôles biologiques cruciaux chez les plantes. Ceux-ci incluent notamment le support mécanique, le transport de l'eau et des nutriments ainsi que la réponse aux stress. Elles fournissent également une abondante source renouvelable d'aliments, de fibres et de bio-carburants. En effet, les parois secondaires représentent la majeure partie de la biomasse végétale. Les graminées font partie des plantes les plus cultivées pour l'alimentation humaine et animale. Ainsi, la compréhension des mécanismes de régulation et de mise en place des parois secondaires chez les graminées est nécessaire pour une exploitation optimale de ces dernières.

L'objectif de ce travail de thèse était de contribuer à la compréhension des mécanismes moléculaires impliqués dans la mise en place des parois cellulaires chez le sorgho. Ses ambitions comprenaient deux volets : l'identification des réseaux de gènes impliqués dans l'établissement des parois et la compréhension fine des modes d'actions des Facteurs de Transcription (FT) contrôlant ces derniers.

Le premier volet a été développé en plusieurs approches. Dans un premier temps, nous nous sommes focalisés sur les FT des familles MYB et NAC qui sont les régulateurs principaux des parois secondaires chez *Arabidopsis*. En effet, peu d'études chez le sorgho ont évalué le rôle des FT responsables de la régulation des parois secondaires alors qu'ils ont le potentiel d'être des régulateurs fins de ce processus. Par leur place centrale dans les réseaux de régulation, la modification de leur expression doit permettre de dévoiler les gènes qu'ils contrôlent. Une analyse phylogénétique de ces deux familles, basée sur cinq espèces de monocotylédones, dont le sorgho, et quatre espèces de dicotylédones, a été mise en place afin d'identifier, grâce au socle de connaissances des FT des parois secondaires chez *Arabidopsis*, le riz, le maïs, le peuplier et l'eucalyptus, les homologues du sorgho potentiellement impliqués dans le même processus.

Dans un deuxième temps, nous avons utilisé des données d'expression pangénomiques décrivant le développement d'entre-nœuds chez le sorgho pour identifier les réseaux de co-expression de gènes impliqués dans ce processus. La construction des réseaux de co-expression repose sur l'agrégation de gènes aux patrons d'expression similaires en faisant l'hypothèse que ces derniers contribuent aux mêmes processus biologiques. L'exploration des réseaux enrichis en gènes et en ontologie géniques (GO-terms) liés à la mise en place des parois a permis d'identifier les groupes de gènes et les FT qui semblent impliqués dans la

mise en place des parois cellulaires. L'utilisation des FT comme gène d'accroche dans les réseaux permet d'identifier les gènes qu'ils apparaissent réguler et peut confirmer ou infirmer l'importance des FT identifiés chez d'autres espèces. L'identification de nouveaux gènes de structure potentiellement cruciaux pour la mise en place des parois est également attendue.

Enfin, nous avons tiré parti des liaisons Phénotype-Génotype identifiées sur la base d'analyses de génétique d'association impliquant des panels de diversité à base génétique large. Nous avons d'abord identifié des zones du génome présentant des associations répétées avec les caractères de composition de la biomasse puis les gènes sous-jacents. La comparaison des listes de gènes issues des analyses de génétique d'association et de réseau de co-expression permet l'accumulation d'indices sur l'implication des gènes dans la mise en place des parois cellulaires.

Le deuxième volet avait pour objectif d'explorer le rôle des FT paraissant être les plus importants dans la mise en place des parois secondaires chez le sorgho. Nous souhaitons donc mettre en place une approche de génétique inverse basée sur l'introduction de mutations aux loci d'intérêt et l'observation des effets phénotypiques induits. Cependant, la difficulté des protocoles de transformation du sorgho nous a amené à nous concentrer sur l'optimisation des méthodes de transformation stable et transitoire.

La transformation génétique stable permet l'introduction de transgènes constitués de séquences mutées ou mutagènes dans du matériel végétal capable de se régénérer en une plante entière et de transmettre la/les mutation(s) à la descendance. La transformation du sorgho se caractérisant par une faible efficacité, l'optimisation des protocoles de culture *in vitro*, de livraison d'ADN transgénique par une méthode biologique et une méthode physique et de sélection des plantes transformées a ainsi été initiée. Dans un premier temps, nous avons ciblé les gènes candidats pour lesquels un effet phénotypique dû à l'intégration du transgène a les meilleures chances d'être observé. Nous avons donc sélectionné des FT du sorgho homologues aux FT impliqués dans la mise en place des parois chez plusieurs autres espèces.

Face à la complexité des essais de transformation stable, nous avons mis en place un protocole de transformation transitoire plus rapide et plus simple sur des protoplastes. L'utilisation de ces cellules isolées a été envisagée dans l'objectif de valider fonctionnellement le rôle de gènes candidats, et notamment les FT, par l'analyse des modifications du transcriptome induites par l'introduction de version mutées. Les protoplastes

peuvent également être mis à profit pour le test de constructions transgéniques avant d'initier des essais de transformation stable avec ceux-ci.

Les résultats attendus pour ce volet étaient la maîtrise de systèmes de transformation stable et transitoire chez le sorgho et des avancées vers la validation du rôle de certains des gènes identifiés dans le cadre de cette étude.

2. Synthèse des résultats

2.1. Analyse de la phylogénie des NAC et des MYB

Les analyses phylogénétiques ont permis d'identifier chez le sorgho 112 FT de la famille des NAC regroupés en 14 clades et 135 FT de la famille des MYB regroupés en 19 clades. Les liens phylogénétiques entre ces FT et leurs homologues de 8 espèces de monocotylédones et de dicotylédones ont été mis à jour.

Vingt-trois FT homologues des gènes de la famille NAC identifiés comme régulateurs des parois secondaires chez d'autres espèces ont été identifiés chez le sorgho. Parmi ceux-ci, 4 homologues des gènes NST/SND, 7 homologues des gènes VND, 5 homologues de NAC075 et 7 homologues VNI2, qui sont des gènes centraux impliqués dans la mise en place des parois cellulaires chez d'autres espèces, ont été mis en évidence. Dix-huit FT MYB ont également été identifiés par leurs liens d'homologies avec les régulateurs des parois secondaires mis en évidence chez d'autres espèces. Il s'agit d'homologues des paires et trio de MYB paralogues MYB83/MYB46, MYB61/MYB55, MYB43/MYB20, MYB58/MYB63, MYB52/MYB54, MYB42/MYB85, MYB4/MYB7/MYB32 et de MYB103 et MYB69 qui agissent chez d'autres espèces à différentes étapes de la régulation de la mise en place des parois. Des clades présentant des diversifications des gènes des monocotylédones par rapport aux dicotylédones ont également été mis en évidence ainsi que des branches de gènes monocotylédones différenciées des gènes dicotylédones.

2.2. Analyse des réseaux de co-expression de gènes

Sur la base de 20 294 gènes exprimés dans les tissus analysés, l'analyse de réseaux de co-expression de gènes a alloué 12 422 gènes à 30 groupes de gènes co-exprimés (appelés modules dans ce document) aux corrélations d'expression supérieures à 0,82. Parmi eux, 6 modules ont été identifiés comme étant enrichis en gènes impliqués dans la mise en place des parois. Les informations qui ont mené à ce résultat sont les patrons d'expression des gènes qu'ils présentent ainsi que les enrichissements en gènes identifiés dans la littérature comme

étant des gènes de biosynthèse des parois et les enrichissements en « GO-terms » liés aux parois. Quatre de ces modules semblent comporter des gènes impliqués dans la mise en place des parois primaires car ils sont enrichis en gènes reliés aux parois primaires comme des gènes de la biosynthèse des pectines et des hémicelluloses et impliqués dans l'élongation cellulaire. Les gènes des deux autres modules semblent, eux, impliqués dans la mise en place des parois secondaires. En effet, ils sont enrichis spécifiquement en GO-terms liés aux phénylpropanoïdes et contiennent des gènes liés aux parois secondaires, comme des CESA spécifiques de la biosynthèse de la cellulose des parois secondaires et des gènes de la voie de biosynthèse des lignines validés chez le sorgho (Bmr). La comparaison des dynamiques d'accumulations des composants pariétaux et des patrons d'expression de ces modules semble être en accord avec cette hypothèse. Il sera intéressant de tester cette cohérence de façon statistique. Les gènes des modules reliés aux parois primaires montrent des pics d'expression précoces dans le développement qui correspondent au stade d'accumulation maximal des polysaccharides de structures et les modules reliés aux parois secondaires atteignent des pics d'expression plus tardifs qui correspondent à l'accumulation des lignines. Les « hub-genes » de ces modules sont majoritairement des gènes de structure dont certains sont annotés avec un ou plusieurs GO-terms liés aux parois. Cependant, un des modules reliés aux parois secondaires possède parmi ses « hub-genes » quatre FT (SbNAC75a, SbMYB52/54, WRKY12, KNAT7) qui sont homologues des régulateurs de la mise en place des parois secondaires chez *Arabidopsis* dont deux sont des FT MYB et NAC.

En plus de l'identification de groupes de gènes potentiellement impliqués dans la mise en place des parois primaires et secondaires, l'analyse des réseaux de co-expression de gènes a permis de mieux caractériser les clades et les fonctions putatives des FT MYB ET NAC.

En effet, nous avons analysé les gènes dont l'expression est la plus corrélée (sous-réseau) avec tous les FT MYB et NAC identifiés chez le sorgho. L'enrichissement en gènes et en GO-terms liés aux parois de ces sous-réseaux nous a permis d'apporter des indices supplémentaires concernant la conservation fonctionnelle de leur rôle au-delà des simples relations d'homologies entre les FT décrits pour leur rôle dans la mise en place des parois secondaires chez différentes espèces et leurs homologues chez le sorgho. Cette analyse nous a également permis de mettre en évidence des FT chez le sorgho dont les sous-réseaux sont enrichis en gènes et GO-terms liés aux parois et qui n'ont aucune relation d'homologie avec des gènes liés à ces fonctions dans d'autres espèces.

Parmi les FT du sorgho homologues à des gènes régulateurs des parois secondaires chez d'autres espèces, nous avons identifié des enrichissements pour au moins un homologue des gènes AtMYB4/AtMYB32/AtMYB7, AtMYB58/AtMYB63, AtMYB52/54, AtMYB20/AtMYB43, AtMYB61, AtMYB55, AtMYB103, AtMYB46/AtMYB83, AtNST1/AtSND1/AtNST2, AtVND6/AtVND7, AtNAC075 et AtVNI2, soient 19 gènes. Vingt FT ne sont pas des homologues proches de gènes impliqués dans les parois secondaires mais présentent tout de même des enrichissements liés aux parois. Parmi ces nouveaux FT, 10 se trouvent dans des branches phylogéniques très éloignées de tout gène annoté comme ayant un rôle dans la régulation des parois secondaires.

2.3.Mise en évidence de régions génomiques associées à la variabilité de la biomasse par analyse de génétique d'association

Cinquante-huit zones regroupant des SNP associés aux caractères liés aux parois, au contenu en sucre et à la digestibilité ont été identifiées en utilisant des panels d'accessions recouvrant la diversité mondiale. Sur une fenêtre de 20kb de part et d'autre des loci d'intérêt, 257 gènes ont été identifiés. Seuls 6 d'entre eux font partie de listes de gènes de structure liés aux parois issues de la littérature. Cent douze gènes font parties des 30 réseaux de co-expression identifiés et 70 gènes sont intégrés dans les 6 réseaux de co-expression reliés aux parois mis en évidence dans cette étude. Plusieurs font partie des gènes co-exprimés avec plusieurs FT identifiés comme régulateurs potentiels de la mise en place des parois dans cette étude. Cependant, aucun de ces FT ne se situe lui-même directement à proximité des loci présentant des associations significatives avec la variabilité de la composition des parois.

2.4.La transformation stable chez le sorgho reste un challenge

Nos essais de transformation par bombardement et *Agrobacterium* ont confirmé l'existence de points critiques lors de la construction des plasmides, du transfert d'ADN et de la régénération des plantes. En effet, le stade de développement des embryons immatures, l'état physiologique des plantes mères, le type d'explant, la concentration de l'antibiotique de sélection, la taille des plasmides ainsi que les promoteurs utilisés sont autant de paramètres qui peuvent avoir un impact très important sur le taux de succès de la transformation. Une seule plante contenant un gène de sélection, *nptII*, est issue des 4000 cals bombardés. La transformation par *Agrobacterium* a permis l'obtention de plantes porteuse du gène *gfp* et d'une construction CRISPR/Cas9, dont l'objectif était d'interrompre la fonction des deux homologues chez le sorgho des gènes NST1/SND1, qui sont des régulateurs principaux de la mise en place des parois secondaires chez *Arabidopsis*. Malheureusement la plante issue de la

cellule transformée ne comportait pas d'édition de la séquence des gènes ciblés et n'a pas transmis la Cas9 à ses descendantes.

2.5. La transformation transitoire sur protoplaste de sorgho est une voie rapide de caractérisation de gènes

Les essais de production et de transformation transitoire de protoplastes ont permis d'optimiser plusieurs points du protocole initial disponible pour le riz. La durée de la digestion enzymatique permettant d'extraire les protoplastes du matériel végétal a été optimisée à 4h dans nos conditions expérimentales et permet d'obtenir jusqu'à 4 millions de protoplastes par gramme de matériel végétal. L'osmolarité du milieu a également été ajustée et permet un taux de viabilité des protoplastes supérieur à 95% et un maintien de 50% de la population initiale 4 jours après l'extraction initiale. L'osmolarité du milieu de transformation a également été optimisée, ce qui nous a permis d'atteindre un taux de transformants supérieur à 70% en routine. Le taux de co-transformation par deux plasmides atteint quant à lui quasiment 100%. Un tel taux de co-transformation est nécessaire pour envisager des essais de validation fonctionnelle. En effet, le principe est d'utiliser le taux de transformation d'un gène marqueur qui, co-transformé à 100% avec un gène d'intérêt, indique directement le taux de protoplastes transformés avec ledit gène d'intérêt.

3. Arabidopsis, un modèle absolu ?

Les résultats obtenus suite aux analyses de génomique comparative et d'analyse de co-expression convergent à suggérer que l'utilisation de plantes modèles comme Arabidopsis est efficace pour l'inférence de fonction à travers l'identification de liens d'homologies.

Comme Zhao et Bartley (2014) et Xiong *et al.* (2005) l'avaient montré, les familles de TF NAC et MYB sont apparues avant la séparation des monocotylédones et des dicotylédones. Nos analyses menées sur 9 espèces différentes confirment ces observations.

Monocotylédones et dicotylédones partagent donc potentiellement de nombreux FT qui trouvent leur origine dans leur ancêtre commun. De plus, comme nous l'avons vu au chapitre I, plusieurs études ont mis en évidence une conservation de la fonction des FT des parois secondaires à travers de nombreuses espèces de plantes supérieures.

Cependant, malgré les similarités observées entre Arabidopsis et les autres espèces, depuis la séparation des monocotylédones et dicotylédones il y a 150 millions d'années, les familles de FT ont divergé de façon disproportionnelle et les réseaux de régulation se sont probablement

différenciés. Les modèles de régulation mis en évidence chez les dicotylédones ne sont donc vraisemblablement pas entièrement généralisables aux monocotylédones (Chaw *et al.*, 2004, Handakumbura et Hazen, 2012), ni à l'ensemble des dicotylédones. L'identification de plusieurs clades de MYB spécifiques aux graminées supporte cette hypothèse (Rabinowicz *et al.*, 1999, Zhao et Bartley, 2014). Les analyses phylogénétiques que nous avons menées sur les familles MYB et NAC ont en outre montré des expansions chez les monocotylédones pour certains clades. Ainsi, les gènes appartenant à ces clades pourraient avoir des rôles spécifiques ou des spécificités spatio-temporelles dans la régulation de la mise en place des parois chez les graminées. A l'inverse, MYB75, qui joue un rôle clé chez *Arabidopsis* dans la bascule entre la formation des parois secondaires et l'accumulation d'anthocyanes, est un exemple de FT spécifique des dicotylédones, dont Zhao et Bartley (2014) n'ont pas retrouvé d'orthologue proche chez les graminées (leur étude n'incluait pas le sorgho).

Des divergences fonctionnelles existent également parmi les acteurs de la régulation des parois secondaires identifiés par homologies entre les dicotylédones et les monocotylédones. Ainsi, les régulateurs NAC du riz (OsSWNs) montrent des spécificités d'expressions et de fonctions, liées à leur localisation dans les vaisseaux ou dans la tige, moins claires que celles observées chez *Arabidopsis* avec les deux groupes NST (fibres) et VND (vaisseaux) (Yoshida *et al.*, 2013). De manière surprenante, même au sein des dicotylédones, la classification des groupes NST et VND est également ambiguë chez le peuplier, une autre dicotylédone (Ohtani *et al.*, 2011, Zhong *et al.*, 2011). Les interactions des gènes NST avec les autres acteurs de la mise en place des parois sont également différentes entre le peuplier et *Arabidopsis* (Li Q. *et al.*, 2012; Lin *et al.*, 2017, Wang *et al.*, 2011).

Chez les MYB, les divergences entre monocotylédones et dicotylédones existent également. Par exemple, ZmMYB11, ZmMYB31 et ZmMYB42, un groupe de FT homologues à AtMYB4 et AtMYB32, régulent négativement certains gènes de la biosynthèse des lignines comme leurs homologues chez *Arabidopsis* (Vélez-Bermúdez *et al.*, 2015). Cependant, l'expression hétérologue de ces 3 ZmMYB chez *Arabidopsis* montre une régulation de gènes différents de ceux mis en évidence chez le maïs (Fornalé *et al.*, 2006, 2010; Sonbol *et al.*, 2009; Vélez-Bermúdez *et al.*, 2015). Une des différences est la régulation de l'enzyme F5H (responsable du ratio S/G des lignines), qui est régulée négativement par ces gènes chez le maïs mais pas lorsqu'ils sont exprimés chez *Arabidopsis* (Sonbol *et al.*, 2009; Fornalé *et al.*, 2010; Vélez-Bermúdez *et al.*, 2015). De telles différences de régulation de la composition en monomères des lignines ainsi que de la régulation de différentes enzymes de la voie de biosynthèse des

lignines ont été observées chez switchgrass et sur feuilles de maïs, riz et sorgho (Shen *et al.*, 2012, Vélez-Bermúdez *et al.*, 2015; Agarwal *et al.*, 2016). Ceci suggère que les graminées pourraient présenter des mécanismes de régulations de la voie de biosynthèse des phénylpropanoïdes différents de ceux existants chez les dicotylédones, en particulier les fonctions des homologues de MYB4 et MYB32 pour la modulation fine des flux de phénylpropanoïdes (Rao et Dixon, 2018).

Malgré un socle commun fort, des divergences de fonctions entre monocotylédones et dicotylédones existent donc. Cependant, sur la base de la revue bibliographique réalisée, aucun gène de régulation de la mise en place des parois secondaires spécifiques aux graminées n'a été caractérisé à ce jour. L'utilisation d'*Arabidopsis* comme socle de connaissances et le relatif faible développement des ressources génétiques et génomiques sur les parois des graminées pourrait être responsable de cet état de fait. En effet, l'existence de différences phylogénétiques et fonctionnelles et de claires divergences de structure et de morphologie entre les dicotylédones et les monocotylédones laisse présager de l'existence de gènes de régulation spécifiques des parois secondaires chez les différents clades de plantes supérieures (Handakumbura et Hazen, 2012).

L'analyse systématique de l'expression de tous les FT de deux familles régulatrices des parois secondaires chez le sorgho a permis de mettre en évidence des FT potentiellement impliqués dans la mise en place des parois qui n'ont pas ou peu de liens avec des régulateurs connus chez *Arabidopsis*. Le travail initié dans cette thèse ouvre donc la voie à la caractérisation d'un réseau de régulation chez le sorgho proche du socle déjà validé chez *Arabidopsis*, le maïs et le riz mais aussi à l'identification de nouveaux FT régulateurs des parois secondaires chez le sorgho qui pourrait révéler l'existence de gènes monocotylédones responsables de régulations spécifiques à cet ordre.

L'utilisation des connaissances acquises chez *Arabidopsis* est donc un outil puissant pour l'inférence fonctionnelle chez les graminées qui peut également servir, en offrant un cadre de comparaison, à la détection de gènes spécifiques qui n'appartiennent pas à des clades communs. Cependant, le processus de mise en place des parois étant complexe, les avancées chez *Arabidopsis* ne peuvent être utilisées qu'en tant que cadre de réflexion pour la caractérisation des réseaux de gènes d'intérêt chez les espèces non modèles. Mais cette approche ne doit pas être exclusive, des efforts de recherche soutenus spécifiques à ces espèces non modèles sont nécessaires.

Dans l'objectif d'identifier des gènes candidats de la mise en place des parois secondaires, des analyses phylogénétiques complémentaires à celles présentées ici peuvent être envisagées. Les FT MYB et NAC ne sont pas les seuls FT impliqués dans la mise en place des PS, les mêmes analyses globales d'exploration des familles KNAT, WRKY ou BLH couplées avec l'exploration des réseaux de gènes impliqués dans la mise en place des parois ont le potentiel de révéler que des orthologues sorgho des gènes connus chez *Arabidopsis* sont aussi impliqués dans la régulation des parois secondaires. Les rôles d'autres gènes de ces familles et d'autres familles d'intérêt, comme WRKY et KNAT dans le cadre de la mise en place des parois, dont la connaissance est uniquement permise par l'analyse phylogénétique, peuvent également potentiellement être révélés par l'analyse des enrichissements en gènes liés aux parois des réseaux de gènes de co-expression.

Les analyses phylogénétiques peuvent également être enrichies d'informations relatives aux pressions évolutives que les séquences subissent. En effet, l'identification de signatures de sélection chez certains membres des familles étudiées comme des variations de fréquences alléliques aux seins de populations globales ou locales est indicatrice de la pression évolutive subie par les gènes pour la fitness des individus. Des gènes dont la séquence et des clades dont les membres subissent une forte pression de sélection ont donc tendance à indiquer des gènes cruciaux pour les fonctions qu'ils occupent. L'identification de tels gènes parmi l'ensemble de candidats permet de focaliser les efforts de validation et de caractérisation des réseaux dont ils font partie pour améliorer notre connaissance des processus dans lesquels ils sont impliqués.

Des analyses de synténies pourraient également être menées. Elles permettent de caractériser la conservation de l'ordre des gènes entre plusieurs espèces apparentées et ainsi d'estimer quels ensembles de gènes sont apparus ou ont disparus d'une espèce à l'autre. Une conservation strictement parallèle de la synténie est un indicateur de la conservation de la fonction desdits gènes entre les espèces tandis qu'une baisse ou une augmentation de synténie souligne une divergence (Bennetzen et Chen, 2008).

4. Découverte et caractérisation de gènes chez les espèces non modèles

La recherche de gènes chez les espèces non modèles peut être favorable lorsque les processus biologiques d'intérêt présentent des divergences d'ampleur, et en particulier s'ils génèrent des qualités agronomiques. Par exemple, chez le sorgho, les parois secondaires sont plus épaisses et présentent un intérêt agronomique bien supérieur à celles d'*Arabidopsis*. La différence d'ampleur est encore plus frappante chez les espèces pérennes ligneuses qui développent un xylème secondaire (bois) très épais. La recherche de gènes candidats chez l'espèce cible devient absolument nécessaire lorsqu'il s'agit de métabolites spécialisés spécifiques à certaines espèces et que les processus n'existent pas chez les plantes modèles. C'est le cas par exemple de la symbiose *Rhizobium*-Fabacée qui est absente chez *Arabidopsis*.

La découverte de gènes dont les fonctions sont absentes chez *Arabidopsis* peut être envisagée par les méthodes développées dans cette thèse quel que soit le processus biologique ciblé. L'inférence de fonction sur la base de liens d'homologies peut être un point de départ permettant d'inscrire quelques gènes connus au sein de leurs familles génétiques. Dans le cas où aucun gène impliqué dans les processus ciblés n'a été identifié chez l'espèce d'intérêt, l'identification des liens d'homologies avec des espèces chez lesquelles le processus a été étudié devient obligatoire.

Cependant, dorénavant chez beaucoup d'espèces, quelques acteurs majeurs au moins de nombreux processus biologiques sont identifiés (que ce soit des gènes de structure ou des FT). L'exploration de réseaux de gènes en se basant sur des gènes d'accroche pour lesquels la fonction est précisément connue peut mener à l'identification de nouveaux gènes impliqués dans les mêmes processus. La systématisation de la démarche à toute une famille de gènes est particulièrement favorable dans des contextes où, comme pour la mise en place des parois secondaires, une grande partie des fonctions est assurée par quelques familles de gènes bien identifiées, ce qui permet d'explorer le réseau de régulation en se focalisant sur un minimum de régulateurs.

En cas d'absence de gène d'accroche ou d'utilisation de gènes aux fonctions inférées par des études phylogénétiques, les approches de réseau de co-expression de gènes chez les espèces non modèles souffrent du manque d'annotations fonctionnelles sur lesquelles baser des études d'enrichissement. En effet, elles sont souvent elles-mêmes inférées à partir de liens d'homologies avec des espèces modèles.

La multiplication des approches est ainsi fondamentale afin d'accumuler des informations aidant à la caractérisation des gènes.

Les approches d'analyse de réseaux, présentées ici et proposées comme un outil d'identification des gènes chez les espèces non-modèles, peuvent être améliorées par la prise en compte de données différentes.

En effet, l'identification de groupes de co-expression est dépendante des contrastes existant dans les données d'expression. Les contrastes dus aux conditions environnementales, aux phases de développement, aux génotypes ou encore aux tissus considérés introduisent de la variation dans les données ce qui permet d'identifier des groupes de gènes dont les variations d'expression sont corrélées entre elles.

Pour l'analyse présentée ici, les contrastes consistaient majoritairement en des stades développementaux différents et des états de régimes hydriques différents. L'analyse plus approfondie des ensembles de gènes impliqués dans la réponse au stress hydrique n'a d'ailleurs pas été spécifiquement couverte et pourrait recevoir un nouvel éclairage en sélectionnant les données de co-expression en fonction des conditions hydriques appliquées. Un niveau de précision encore supérieur pourrait être atteint en collectant des données d'expression sur des pas de temps plus fins, à proximité de l'application des stress hydriques et de la récupération après irrigation, et faire apparaître des groupes de gènes impliqués dans la réponse au stress hydrique.

Par ailleurs, la prise de données sur des pas de temps fins pourrait améliorer l'identification des groupes de co-expression quel que soit le contraste de conditions considéré. Ainsi, il serait possible de tester l'effet sur l'identification de réseaux de co-expression des contrastes résultant de différents génotypes, de phénotypes contrastés pour les caractères d'intérêt (composition des parois riches en sucre ou au contraire en lignine, etc.), ou de l'analyse des différentes couches cellulaires verticales d'âge différents d'un entre-nœuds suivi au cours de son développement. En effet, les types cellulaires sont très variées au sein de l'entre-nœud et les caractéristiques des parois qui les entourent varient que ce soient entre les types de tissus ou les âges de tissus qui cohabitent sur la hauteur de l'entre-nœud (partie basse plus jeune que la partie haute). Même si l'analyse du transcriptome de l'ensemble des cellules d'un entre-nœud au cours de son développement permet une vision globale des variations d'expression, la prise en compte des tissus voire des types cellulaires permettrait d'apporter des nouvelles informations relatives aux différences de groupes de gènes impliqués. Pour pouvoir mener ce

genre d'analyses, la microdissection laser sur groupes de cellules ou cellule individuelle (LCM) se développe. Elle permet d'accéder aux transcriptomes de cellules isolées ou de groupes de cellules spécifiques. Il serait ainsi possible d'envisager d'isoler des cellules qui sont dépourvues de parois secondaires et d'autres qui en possèdent tout en séparant les cellules de soutien des vaisseaux ou les cellules du pourtour de la tige et celles en son cœur qui ont des caractéristiques de parois très différentes. Enfin, la construction de réseaux de co-expression à partir de données de transcriptomiques générées par d'autres études et non valorisées par ces approches pourrait offrir une base de comparaison avec le jeu de donnée utilisé dans cette étude. Globalement, la comparaison des groupes de gènes co-exprimés construites à partir de données transcriptomiques collectées dans des conditions et des tissus différents permettrait de vérifier le caractère générique de nos conclusions.

Les groupes de co-expression de gènes mis en évidence l'ont été en partie grâce à la connaissance, issue de la littérature, d'un certain nombre de gènes de structures impliqués dans la biosynthèse des composés pariétaux. La séparation entre ceux potentiellement impliqués dans les parois primaires et ceux impliqués dans les parois secondaires est intervenue majoritairement grâce à la comparaison entre les patrons d'expression et les dynamiques d'accumulation des composés pariétaux en prenant la lignine comme indicateur des parois secondaires. La connaissance de certains gènes clés des parois secondaires (Bmr, CESA) a permis de consolider cette analyse. Si l'identification de ces 2 types de groupes de co-expression (parois primaires et secondaires) permet de subdiviser la liste de gènes connus en deux catégories selon qu'ils semblent impliqués dans les parois primaires ou secondaires, l'identification plus fine des gènes de structures liés aux différentes catégories de polysaccharides comme par exemple les MLG ou les XyG, qui sont spécifiques des parois primaires chez les graminées, permettrait de conforter d'autant plus nos conclusions. Ceci pourrait être réalisé, à nouveau, par une comparaison avec les gènes de structure identifiés chez les autres espèces. De surcroît, dans le cadre de nos études, les gènes de structures des parois ont majoritairement été utilisés comme des outils pour caractériser l'enrichissement des groupes de co-expression. Ils pourraient être mis à profit comme des gènes d'accroche afin d'interroger les réseaux dont ils font partie et identifier leurs plus proches régulateurs potentiels.

Chez les espèces non modèles, la découverte de gènes candidats se heurte à de nombreux freins concernant la disponibilité et la précision des ressources génétiques et génomiques. En effet, les approches de génomique comparative et de réseaux de gènes dépendent de la

disponibilité d'un génome séquencé, dont les gènes sont annotés (motifs fonctionnels, début et fin de séquence, etc.) et des variants d'épissage identifiés, ainsi que des informations concernant la fonction putative des gènes, comme les annotations automatiques basées sur les homologues chez les espèces modèles ou l'association à des GO-terms fonctionnels. Que ce soit chez les espèces modèles ou non modèles, les approches de bioinformatique doivent être complétées par d'autres études, en particulier fonctionnelles. Les autres méthodes de recherche de gènes candidats abordées dans ce travail chez le sorgho reposent sur des analyses d'association Phénotype-Génotype à grande échelle et de génétique inverse qui souffrent en partie des mêmes limitations que celles évoquées ci-dessus. En plus de celles-ci, ces approches nécessitent la disponibilité de plateformes de génotypage d'individus et de grandes populations, d'une analyse du phénotype précise et haut-débit, de la disponibilité de banques de mutants, et de protocoles de transformation efficaces.

5. Les difficultés de la transformation génétique

Si nos essais de transformation transitoire sur des protoplastes de sorgho fonctionnent remarquablement bien, la transformation stable est encore un challenge chez cette espèce. En tant qu'outil de validation du rôle des gènes, la majorité des espèces a déjà fait l'objet d'essais de transformation. Chez toutes les plantes, les principales limitations de la transformation stable sont la livraison d'ADN et la régénération de plantes entières à partir de cellules totipotentes. Ces deux problématiques ne sont pas indépendantes car il est souvent difficile de combiner totipotence et compétence à la transformation (Sharma *et al.*, 2004). C'est en particulier le cas pour la méthode la plus répandue de transformation qui implique le vecteur biologique *Agrobacterium*. Si chez certaines espèces, comme le tabac, la totipotence et la compétence à la transformation sont aisément atteignables et compatibles, la transformation est laborieuse chez les monocotylédones, les gymnospermes et les fabacées qui ne sont pas naturellement susceptibles à l'infection par *Agrobacterium*. L'impossibilité des cellules totipotentes à être aptes à la transformation peut être surmontée par des voies qui s'affranchissent des contraintes biologiques, comme la biolistique, qui est théoriquement capable de transformer n'importe quel type de cellule totipotente. Cette voie peut d'ailleurs également être utilisée pour des essais de transformation transitoire sur des tissus d'intérêt quand la régénération n'est pas acquise. Une autre voie de contournement, qui a obtenu des succès chez *Medicago truncatula*, une fabacée, et qui permet, elle, de s'affranchir de la régénération, consiste en l'application directe d'*Agrobacterium* sur des fleurs ou des jeunes plantules, en favorisant ou non sa pénétration par l'application de vide (Trieu *et al.*, 2001). La

multiplication des techniques est un indicateur des difficultés existantes pour la transformation stable chez toutes les espèces de plantes et des efforts des chercheurs pour les surmonter.

Les objectifs de la transformation stable chez le sorgho, une fois qu'elle sera maîtrisée, est de permettre la validation fonctionnelle des gènes candidats. En particulier, dans le cadre de la mise en place des parois secondaires, la validation du rôle des FT identifiés et de ceux dont aucun homologue chez d'autres espèces n'a de rôle annoté dans ce processus sera primordiale. Bien que la communauté internationale en tant qu'ensemble soit capable de générer des mutants de sorgho, la maîtrise des protocoles de transformation au niveau des laboratoires est laborieuse et longue à atteindre. La mise en place des protocoles au niveau local s'impose pour garantir l'indépendance de la recherche, cependant, la longueur de ces essais pose la question de la transmission des techniques entre laboratoires au-delà de la simple référence à des protocoles publiés. Une mission de transfert de technologie a eu lieu au cours de cette thèse dans ce sens mais n'a malheureusement pas permis d'aboutir à l'obtention de mutants. En prenant en compte la sensibilité des protocoles, la sous-traitance directe de ces essais chez des experts à l'international est probablement la solution la plus efficace d'obtention de plantes transformées aujourd'hui. Les difficultés de développement des méthodes de transformation génétique chez les espèces non modèles font apparaître comme obligatoire le recours à d'autres ressources comme celles existant chez les espèces modèles. En particulier, et malgré les divergences déjà évoquées, l'étude des processus qui sont potentiellement bien conservés en raison de leur importance dans des fonctions communes à toutes les plantes, comme la mise en place des parois, peut bénéficier de l'utilisation des ressources disponibles chez *Arabidopsis* et les autres plantes modèles.

Ainsi, même si la transformation du sorgho n'est pas encore maîtrisée au laboratoire, il est tout de même envisageable de tester le rôle des gènes candidats identifiés dans le cadre d'essais d'expression hétérologues. En effet, les études de complémentation de phénotypes mutants chez des espèces qui peuvent être facilement transformées, comme *Arabidopsis* et le riz, par des gènes homologues d'autres espèces, ici le sorgho, permet d'avoir une idée du niveau de conservation de la fonction de ces gènes entre les espèces. Dans la même logique, le rôle des gènes identifiés chez le sorgho peut être mis en évidence chez des espèces proches. Ainsi, les gènes candidats identifiés par ce travail sont en cours d'évaluation chez le riz par des essais de transformation avec vecteur CRISPR/Cas9 ciblant les homologues de ces gènes chez le riz. De plus, un pas vers la caractérisation fonctionnelle des gènes d'intérêt pourrait

être franchi en utilisant des approches qui dépendent moins de l'espèce et du génotype que la transformation, comme l'hybridation *in situ*, qui permettrait l'analyse de la localisation dans les différents tissus des transcrits des gènes identifiés. En particulier, de telles approches réalisées sur des coupes de tiges de sorgho permettraient de conclure sur la spécificité de gènes comme VND7 et en général les gènes de la famille VND qui, chez *Arabidopsis*, sont spécifiques des tissus vasculaires tandis que chez d'autres espèces, comme le peuplier et eucalyptus, la spécificité d'expression est moins claire. Chez le sorgho, les informations collectées sur le rôle de ces gènes suggèrent également des différences par rapport à *Arabidopsis*.

L'optimisation de la transformation de protoplastes de sorgho permet de disposer d'une technique de transition entre l'identification de gènes candidats par des approches globales et la validation de leur rôle via la transformation stable, mais permet aussi l'acquisition d'informations fonctionnelles sur le rôle de ces gènes. Ainsi, les constructions génétiques destinées à être utilisées pour caractériser le rôle des gènes candidats en transformation stable peuvent être en premier lieu testées sur des protoplastes et leur activité évaluée par du génotypage (en cas de test de constructions KO CRISPR/Cas9) ou l'étude de l'expression des gènes transformés et de quelques-unes de ses cibles (en cas de test de constructions de surexpression). Les constructions envisagées pour tester le rôle de gènes candidats en transformation transitoire peuvent également se diversifier par exemple par l'utilisation de promoteurs inductibles, qui permettent d'induire sur commande l'expression des gènes afin de détecter les gènes en aval dont l'expression est modifiée par l'activité du transgène. Cette approche permet de caractériser précisément la dynamique d'expression des cibles et de la voie complète. L'analyse globale du transcriptome couplée à l'utilisation de ce genre de construction s'inscrit dans une démarche de caractérisation du rôle des gènes au-delà du simple test des constructions. Les autres constructions présentées (CRISPR/Cas9, surexpression) pourraient ainsi, elles aussi, être utilisées dans le cadre de la caractérisation fonctionnelle en comparant le transcriptome de protoplastes transformés et celui de protoplastes non transformés. La caractérisation du rôle de gènes impliqués dans la mise en place des parois secondaires est cependant difficilement envisageable au-delà de la transcriptomique en raison de l'absence de paroi des protoplastes. Cependant, les modifications visibles du phénotype induites par le transgène pourraient être observées sur des protoplastes âgés d'une semaine ou plus, puisque la paroi commence à être reconstruite dès 7 jours. La caractérisation du rôle de gènes impliqués dans les nombreux processus actifs chez

les protoplastes pourrait par contre être envisagée selon les effets attendus sur le métabolisme qui reste observable chez les protoplastes.

Enfin, la régénération des protoplastes transformés en plantes entières est une perspective à long terme de maîtrise de la transformation stable. Cependant, malgré la maîtrise de la transformation efficace de protoplastes, la régénération *in vitro* et la sélection des plantules transformées restent des étapes clés du processus d'obtention de plantes transformées de façon stable. La régénération de cals de sorgho issus d'embryons immatures semble possible relativement facilement en culture *in vitro* mais la modification du type d'explant (protoplastes contre embryons immatures) peut entraîner l'échec de la régénération.

Contributions

1. Contributions au chapitre II

Les essais de transcriptomique 2013, 2014 et 2015 décrits dans cette section se sont déroulés avant mon arrivée en thèse, je n'ai donc pas participé à l'échantillonnage ni à la construction des banques RNAseq.

Les essais de génétique d'association ont également eu lieu avant le début de ma thèse et les analyses ont été effectuées par I. Vilnius. Seule la récupération de zones d'intérêts, des identifiants de gènes et leur utilisation pour d'autres analyses ont été faites dans le cadre de cette thèse.

2. Contributions au chapitre III

Le chapitre III porte sur plusieurs expériences de transformation du sorgho. Il résulte du travail de plusieurs personnes.

Lauriane Hennen (LH), Maëlle Rios (MR) et Nancy Terrier (NT) ont cloné les constructions ZmUbi ::SbMYB60b, sMYB46, ZmUbi ::VND7b. MR a cloné les constructions mNST et 35S::VND7b.

LH, Frédérique Richaud (FR), MR et NT ont participé aux transformations par biolistique. En particulier LH sur toutes les étapes de la 1ère, l'extraction des embryons et sur le bombardement des suivantes, FR, MR et NT sur l'extraction des embryons et FR et MR majoritairement pour le suivi des repiquages et la préparation des milieux. LH et MR ont effectué les génotypages.

Caroline Calatayud (CC), LH, FR, MR et NT ont participé aux transformations par *Agrobacterium*. En particulier, CC, FR et NT aux extractions d'embryons, CC et FR aux repiquages et préparation des milieux. CC a effectué les génotypages.

Angélique Berger (AB), CC, LH, FR et NT ont participé aux transformations de protoplastes. AB, CC et FR ont majoritairement mené les expérimentations.

Références

- Adli, Mazhar. 2018. "The CRISPR Tool Kit for Genome Editing and Beyond." *Nature Communications* 9 (1): 1911. <https://doi.org/10.1038/s41467-018-04252-2>.
- Afgan, Enis, Dannon Baker, Bérénice Batut, Marius Van Den Beek, Dave Bouvier, Martin Ech, John Chilton, et al. 2018. "The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2018 Update." *Nucleic Acids Research* 46 (W1): W537–44. <https://doi.org/10.1093/nar/gky379>.
- Agarwal, Tina, Erich Grotewold, Andrea I Doseff, and John Gray. 2016. "MYB31 / MYB42 Syntelogs Exhibit Divergent Regulation of Phenylpropanoid Genes in Maize , Sorghum and Rice." *Nature Publishing Group*, no. May: 1–17. <https://doi.org/10.1038/srep28502>.
- Agati, Giovanni, Elisa Azzarello, Susanna Pollastri, and Massimiliano Tattini. 2012. "Flavonoids as Antioxidants in Plants: Location and Functional Significance." *Plant Science* 196 (November): 67–76. <https://doi.org/10.1016/j.plantsci.2012.07.014>.
- Aharoni, Asaph, Shital Dixit, Reinhard Jetter, Eveline Thoenes, Gert van Arkel, and Andy Pereira. 2004. "The SHINE Clade of AP2 Domain Transcription Factors Activates Wax Biosynthesis, Alters Cuticle Properties, and Confers Drought Tolerance When Overexpressed in Arabidopsis." *The Plant Cell* 16 (9): 2463–80. <https://doi.org/10.1105/tpc.104.022897>.
- Ahmed, Rana Imtiaz, Anming Ding, Minmin Xie, and Yingzhen Kong. 2018. "Progress in Optimization of Agrobacterium-Mediated Transformation in Sorghum (Sorghum Bicolor)." *International Journal of Molecular Sciences* 19 (10). <https://doi.org/10.3390/ijms19102983>.
- Aida, M, T Ishida, H Fukaki, H Fujisawa, and M Tasaka. 1997. "Genes Involved in Organ Separation in Arabidopsis: An Analysis of the Cup-Shaped Cotyledon Mutant." *The Plant Cell* 9 (6): 841–57. <https://doi.org/10.1105/tpc.9.6.841>.
- Albenne, Cécile, Hervé Canut, Georges Boudart, Yu Zhang, Hélène San Clemente, Rafael Pont-Lezica, and Elisabeth Jamet. 2009. "Plant Cell Wall Proteomics: Mass Spectrometry Data, a Trove for Research on Protein Structure/Function Relationships." *Molecular Plant* 2 (5): 977–89. <https://doi.org/10.1093/mp/ssp059>.
- Albenne, Cécile, Hervé Canut, Laurent Hoffmann, and Elisabeth Jamet. 2014. "Plant Cell Wall Proteins: A Large Body of Data, but What about Runaways?" *Proteomes* 2 (2): 224–42. <https://doi.org/10.3390/proteomes2020224>.
- Albersheim P, Darvill A, Roberts K, Sederoff R, Staehelin A (2011) "Plant Cell Walls, from Chemistry to Biology." Garland Science, New York, pp 227–272
- Alexa A., Rahnenfuhrer J. (2019). topGO: Enrichment Analysis for Gene Ontology. R package version 2.38.1. Available at: <https://www.bioconductor.org/packages/release/bioc/html/topGO.html> (accessed February 20, 2020).
- Allan, Andrew C., Roger P. Hellens, and William A. Laing. 2008. "MYB Transcription Factors That Colour Our Fruit." *Trends in Plant Science* 13 (3): 99–102.

- <https://doi.org/10.1016/j.tplants.2007.11.012>.
- Almodares, A., and M. R. Hadi. 2009. "Production of Bioethanol from Sweet Sorghum: A Review." *African Journal of Agricultural Research* 4 (9): 772–80.
- Altpeter, Fredy, Nathan M. Springer, Laura E. Bartley, Ann E. Blechl, Thomas P. Brutnell, Vitaly Citovsky, Liza J. Conrad, et al. 2016. "Advancing Crop Transformation in the Era of Genome Editing." *Plant Cell* 28 (7): 1510–20. <https://doi.org/10.1105/tpc.16.00196>.
- Amandeep Kaur, Saroj Kumar Sah. 2014. "Genetic Transformation of Rice: Problems, Progress and Prospects." *Rice Research: Open Access* 03 (01): 1–10. <https://doi.org/10.4172/2375-4338.1000132>.
- An, P., X. Li, Y. Zheng, A. Matsuura, J. Abe, A. E. Eneji, E. Tanimoto, and S. Inanaga. 2014. "Effects of NaCl on Root Growth and Cell Wall Composition of Two Soya Bean Cultivars with Contrasting Salt Tolerance." *Journal of Agronomy and Crop Science* 200 (3): 212–18. <https://doi.org/10.1111/jac.12060>.
- Anami, Sylvester, Elizabeth Njuguna, Griet Coussens, Stijn Aesaert, and Mieke Van Lijsebettens. 2013. "Higher Plant Transformation: Principles and Molecular Tools." *International Journal of Developmental Biology* 57 (6–8): 483–94. <https://doi.org/10.1387/ijdb.130232mv>.
- Andersen, M. C. F. (2014). "Synthesis and Application of Plant Cell Wall Oligogalactans." Department of Chemistry, Technical University of Denmark
- Andrews S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (accessed February 20, 2020).
- Armstrong, CharlesL., WilliamL. Petersen, WallaceG. Buchholz, BenjaminA. Bowen, and SallyL. Sulc. 1990. "Factors Affecting PEG-Mediated Stable Transformation of Maize Protoplasts." *Plant Cell Reports* 9 (6). <https://doi.org/10.1007/BF00232864>.
- Asai, Tsuneaki, Guillaume Tena, Joulia Plotnikova, Matthew R. Willmann, Wan-Ling Chiu, Lourdes Gomez-Gomez, Thomas Boller, Frederick M. Ausubel, and Jen Sheen. 2002. "MAP Kinase Signalling Cascade in Arabidopsis Innate Immunity." *Nature* 415 (6875): 977–83. <https://doi.org/10.1038/415977a>.
- Aufrère, J., R. Baumont, L. Delaby, J. R. Peccatte, J. Andrieu, J. P. Andrieu, and J. P. Dulphy. 2007. "Prévision de La Digestibilité Des Fourrages Par La Méthode Pepsine-Cellulase. Le Point Sur Les Équations Proposées." *Productions Animales* 20 (2): 129–36.
- Bae, Sangsu, Jiyeon Kweon, Heon Seok Kim, and Jin-Soo Kim. 2014. "Microhomology-Based Choice of Cas9 Nuclease Target Sites." *Nature Methods* 11 (7): 705–6. <https://doi.org/10.1038/nmeth.3015>.
- Balazadeh, Salma, Hamad Siddiqui, Annapurna D. Allu, Lilian P. Matallana-Ramirez, Camila Caldana, Mohammad Mehrnia, Maria-Inés Zanol, Barbara Köhler, and Bernd Mueller-Roeber. 2010. "A Gene Regulatory Network Controlled by the NAC Transcription Factor ANAC092/AtNAC2/ORE1 during Salt-Promoted Senescence." *The Plant Journal* 62 (2): 250–

64. <https://doi.org/10.1111/j.1365-313X.2010.04151.x>.
- Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. 2007. "CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes." *Science* 315 (5819): 1709–12. <https://doi.org/10.1126/science.1138140>.
- Barros, Jaime, Luis Escamilla-Trevino, Luhua Song, Xiaolan Rao, Juan Carlos Serrani-Yarce, Maite Docampo Palacios, Nancy Engle, et al. 2019. "4-Coumarate 3-Hydroxylase in the Lignin Biosynthesis Pathway Is a Cytosolic Ascorbate Peroxidase." *Nature Communications* 10 (1): 1–11. <https://doi.org/10.1038/s41467-019-10082-7>.
- Barros, Jaime, Henrik Serk, Irene Granlund, and Edouard Pesquet. 2015. "The Cell Biology of Lignification in Higher Plants." *Annals of Botany* 115 (7): 1053–74. <https://doi.org/10.1093/aob/mcv046>.
- Bashline, Logan, Lei Lei, Shundai Li, and Ying Gu. 2014. "Cell Wall, Cytoskeleton, and Cell Expansion in Higher Plants." *Molecular Plant* 7 (4): 586–600. <https://doi.org/10.1093/mp/ssu018>.
- Battraw, M., and T. C. Hall. 1991. "Stable Transformation of Sorghum Bicolor Protoplasts with Chimeric Neomycin Phosphotransferase II and β -Glucuronidase Genes." *Theoretical and Applied Genetics* 82 (2): 161–68. <https://doi.org/10.1007/BF00226207>.
- Beckman, Carl H. 2000. "Phenolic-Storing Cells: Keys to Programmed Cell Death and Periderm Formation in Wilt Disease Resistance and in General Defence Responses in Plants?" *Physiological and Molecular Plant Pathology* 57 (3): 101–10. <https://doi.org/10.1006/pmpp.2000.0287>.
- Belide, Srinivas, Thomas Vanhercke, James Robertson Petrie, and Surinder Pal Singh. 2017. "Robust Genetic Transformation of Sorghum (Sorghum Bicolor L.) Using Differentiating Embryogenic Callus Induced from Immature Embryos." *Plant Methods* 13 (1): 1–12. <https://doi.org/10.1186/s13007-017-0260-9>.
- Bennetzen, Jeffrey L., and Mingsheng Chen. 2008. "Grass Genomic Synteny Illuminates Plant Genome Function and Evolution." *Rice* 1 (2): 109–18. <https://doi.org/10.1007/s12284-008-9015-6>.
- Berthet, Serge, Nathalie Demont-Caulet, Brigitte Pollet, Przemyslaw Bidzinski, Laurent Cézard, Phillipe Le Bris, Nero Borrega, et al. 2011. "Disruption of LACCASE4 and 17 Results in Tissue-Specific Alterations to Lignification of Arabidopsis Thaliana Stems." *The Plant Cell* 23 (3): 1124–37. <https://doi.org/10.1105/tpc.110.082792>.
- Bhargava, Apurva, Shawn D. Mansfield, Hardy C. Hall, Carl J. Douglas, and Brian E. Ellis. 2010. "MYB75 Functions in Regulation of Secondary Cell Wall Formation in the Arabidopsis Inflorescence Stem." *Plant Physiology* 154 (3): 1428–38. <https://doi.org/10.1104/pp.110.162735>.
- Bhatia, Rakesh, Sue Dalton, Luned A. Roberts, Odin M. Moron-Garcia, Rosario Iacono, Ondrej Kosik, Joe A. Gallagher, and Maurice Bosch. 2019. "Modified Expression of ZmMYB167 in

- Brachypodium Distachyon and Zea Mays Leads to Increased Cell Wall Lignin and Phenolic Content.” *Scientific Reports* 9 (1): 8800. <https://doi.org/10.1038/s41598-019-45225-9>.
- Bhojwani S.S., Razdan, M.K. 1996. *Plant Tissue Culture : Plant Tissue Culture* : Elsevier Science.
- Bhuiyan, N. H., G. Selvaraj, Y. Wei, and J. King. 2009. “Gene Expression Profiling and Silencing Reveal That Monolignol Biosynthesis Plays a Critical Role in Penetration Defence in Wheat against Powdery Mildew Invasion.” *Journal of Experimental Botany* 60 (2): 509–21. <https://doi.org/10.1093/jxb/ern290>.
- Bibikova, M., D. Carroll, D. J. Segal, J. K. Trautman, J. Smith, Y.-G. Kim, and S. Chandrasegaran. 2001. “Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases.” *Molecular and Cellular Biology* 21 (1): 289–97. <https://doi.org/10.1128/MCB.21.1.289-297.2001>.
- Billot, Claire, Punna Ramu, Sophie Bouchet, Jacques Chantreau, Monique Deu, Laetitia Gardes, Jean Louis Noyer, et al. 2013. “Massive Sorghum Collection Genotyped with SSR Markers to Enhance Use of Global Genetic Resources.” *PLoS ONE* 8 (4): 1–16. <https://doi.org/10.1371/journal.pone.0059714>.
- Biswal, Ajaya K, Zhangying Hao, Sivakumar Pattathil, Xiaohan Yang, Kim Winkeler, Cassandra Collins, Sushree S Mohanty, et al. 2015. “Downregulation of GAUT12 in Populus Deltoides by RNA Silencing Results in Reduced Recalcitrance, Increased Growth and Reduced Xylan and Pectin in a Woody Biofuel Feedstock.” *Biotechnology for Biofuels* 8 (1): 41. <https://doi.org/10.1186/s13068-015-0218-y>.
- Block, Marc De, Dirk De Brouwer, and Paul Tenning. 1989. “Transformation of Brassica Napus and Brassica Oleracea Using Agrobacterium Tumefaciens and the Expression of the Bar and Neo Genes in the Transgenic Plants.” *Plant Physiology* 91 (2): 694–701. <https://doi.org/10.1104/pp.91.2.694>.
- Block, Marc De, Luis Herrera-estrella, Marc Van Montagu, Jeff Schell, and Patricia Zambryski. 1984. “In Regenerated Plants And” 3 (8): 1681–89. <https://doi.org/10.1002/j.1460-2075.1984.tb02032.x>.
- Boch, J., H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt, and U. Bonas. 2009. “Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors.” *Science* 326 (5959): 1509–12. <https://doi.org/10.1126/science.1178811>.
- Boerjan, Wout, John Ralph, and Marie Baucher. 2003. “Lignin Biosynthesis.” *Annual Review of Plant Biology* 54 (1): 519–46. <https://doi.org/10.1146/annurev.arplant.54.031902.134938>.
- Bonawitz, Nicholas D., and Clint Chapple. 2010. “The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype.” *Annual Review of Genetics* 44 (1): 337–63. <https://doi.org/10.1146/annurev-genet-102209-163508>.
- Bosch, Maurice, Claus-Dieter Mayer, Alan Cookson, and Iain S. Donnison. 2011. “Identification of Genes Involved in Cell Wall Biogenesis in Grasses by Differential Gene Expression Profiling of

- Elongating and Non-Elongating Maize Internodes.” *Journal of Experimental Botany* 62 (10): 3545–61. <https://doi.org/10.1093/jxb/err045>.
- Brandt, Agnieszka, John Gräsvik, Jason P. Hallett, and Tom Welton. 2013. “Deconstruction of Lignocellulosic Biomass with Ionic Liquids.” *Green Chemistry* 15 (3): 550. <https://doi.org/10.1039/c2gc36364j>.
- Brenton, Zachary W, Elizabeth A Cooper, Mathew T Myers, Richard E Boyles, Nadia Shakoor, Kelsey J Zielinski, Bradley L Rauh, William C Bridges, Geoffrey P Morris, and Stephen Kresovich. 2016. “A Genomic Resource for the Development, Improvement, and Exploitation of Sorghum for Bioenergy The Designation of a Functional Model Grass Species and The” 204 (September): 21–33. <https://doi.org/10.1534/genetics.115.183947>.
- Brett, C., and K. Waldron. 1990. “The Cell Wall and Intercellular Transport.” In *Physiology and Biochemistry of Plant Cell Walls*, 114–36. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-9641-6_5.
- Brown, David M., Zhinong Zhang, Elaine Stephens, Paul Dupree, and Simon R. Turner. 2009. “Characterization of IRX10 and IRX10-like Reveals an Essential Role in Glucuronoxylan Biosynthesis in Arabidopsis.” *The Plant Journal* 57 (4): 732–46. <https://doi.org/10.1111/j.1365-313X.2008.03729.x>.
- Brown, R. Malcolm. 1996. “The Biosynthesis of Cellulose.” *Journal of Macromolecular Science, Part A* 33 (10): 1345–73. <https://doi.org/10.1080/10601329608014912>.
- Burks, Payne S., Chris M. Kaiser, Elizabeth M. Hawkins, and Patrick J. Brown. 2015. “Genomewide Association for Sugar Yield in Sweet Sorghum.” *Crop Science* 55 (5): 2138. <https://doi.org/10.2135/cropsci2015.01.0057>.
- Burr, Sally J., and Stephen C. Fry. 2009. “Extracellular Cross-Linking of Maize Arabinoxylans by Oxidation of Feruloyl Esters to Form Oligoferuloyl Esters and Ether-like Bonds.” *The Plant Journal* 58 (4): 554–67. <https://doi.org/10.1111/j.1365-313X.2009.03800.x>.
- Burris, Kellie P., Elizabeth M. Dlugosz, A. Grace Collins, C. Neal Stewart, and Scott C. Lenaghan. 2016. “Development of a Rapid, Low-Cost Protoplast Transfection System for Switchgrass (*Panicum Virgatum* L.).” *Plant Cell Reports* 35 (3): 693–704. <https://doi.org/10.1007/s00299-015-1913-7>.
- Burton, R. A. 2006. “Cellulose Synthase-Like CslF Genes Mediate the Synthesis of Cell Wall (1,3;1,4)-D-Glucans.” *Science* 311 (5769): 1940–42. <https://doi.org/10.1126/science.1122975>.
- Burton, Rachel A., Michael J. Gidley, and Geoffrey B. Fincher. 2010. “Heterogeneity in the Chemistry, Structure and Function of Plant Cell Walls.” *Nature Chemical Biology* 6 (10): 724–32. <https://doi.org/10.1038/nchembio.439>.
- Busse-Wicher, Marta, Thiago C. F. Gomes, Theodora Tryfona, Nino Nikolovski, Katherine Stott, Nicholas J. Grantham, David N. Bolam, Munir S. Skaf, and Paul Dupree. 2014. “The Pattern of Xylan Acetylation Suggests Xylan May Interact with Cellulose Microfibrils as a Twofold Helical

- Screw in the Secondary Plant Cell Wall of *Arabidopsis Thaliana*.” *The Plant Journal* 79 (3): 492–506. <https://doi.org/10.1111/tpj.12575>.
- Buxton, Dwayne R., and Daren D. Redfearn. 1997. “Plant Limitations to Fiber Digestion and Utilization.” *The Journal of Nutrition* 127 (5): 814S–818S. <https://doi.org/10.1093/jn/127.5.814S>.
- Caffall, Kerry Hosmer, and Debra Mohnen. 2009. “The Structure, Function, and Biosynthesis of Plant Cell Wall Pectic Polysaccharides.” *Carbohydrate Research* 344 (14): 1879–1900. <https://doi.org/10.1016/j.carres.2009.05.021>.
- Cai, Bin, Cheng Hui Li, and Jian Huang. 2014. “Systematic Identification of Cell-Wall Related Genes in *Populus* Based on Analysis of Functional Modules in Co-Expression Network.” *PLoS ONE* 9 (4). <https://doi.org/10.1371/journal.pone.0095176>.
- Cai, Zhuo, Hiroshi Nakajima, Margaret Woldu, Anders Berglund, Maud Bergman, and Toru Okabe. 1999. “In Vitro Corrosion Resistance of Titanium Made Using Different Fabrication Methods.” *Biomaterials* 20 (2): 183–90. [https://doi.org/10.1016/S0142-9612\(98\)00160-4](https://doi.org/10.1016/S0142-9612(98)00160-4).
- Cantu, Dario, Ariel R. Vicente, John M. Labavitch, Alan B. Bennett, and Ann L.T. Powell. 2008. “Strangers in the Matrix: Plant Cell Walls and Pathogen Susceptibility.” *Trends in Plant Science* 13 (11): 610–17. <https://doi.org/10.1016/j.tplants.2008.09.002>.
- Cao, Jianmei, Dongmei Yao, Fan Lin, and Mingyi Jiang. 2014. “PEG-Mediated Transient Gene Expression and Silencing System in Maize Mesophyll Protoplasts: A Valuable Tool for Signal Transduction Study in Maize.” *Acta Physiologiae Plantarum* 36 (5): 1271–81. <https://doi.org/10.1007/s11738-014-1508-x>.
- Capella-Gutiérrez, Salvador, José M. Silla-Martínez, and Toni Gabaldón. 2009. “TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses.” *Bioinformatics* 25 (15): 1972–73. <https://doi.org/10.1093/bioinformatics/btp348>.
- Carpita, Nicholas C. 1996. “Structure and biogenesis of the cell walls of grasses.” *Annual Review of Plant Physiology and Plant Molecular Biology* 47 (1): 445–76. <https://doi.org/10.1146/annurev.arplant.47.1.445>.
- Carpita, Nicholas C. 2011. “Update on Mechanisms of Plant Cell Wall Biosynthesis: How Plants Make Cellulose and Other (1→4)-β-D-Glycans.” *Plant Physiology* 155 (1): 171–84. <https://doi.org/10.1104/pp.110.163360>.
- Carvalho, Carlos Henrique S., Usha B. Zehr, Nilupa Gunaratna, Joseph Anderson, Halina H. Kononowicz, Thomas K. Hodges, and John D. Axtell. 2004. “Agrobacterium-Mediated Transformation of Sorghum: Factors That Affect Transformation Efficiency.” *Genetics and Molecular Biology* 27 (2): 259–69. <https://doi.org/10.1590/S1415-47572004000200022>.
- Casas, A M, A K Kononowicz, U B Zehr, D T Tomes, J D Axtell, L G Butler, R A Bressan, and P M Hasegawa. 1993. “Transgenic Sorghum Plants via Microprojectile Bombardment.” *Proceedings of the National Academy of Sciences* 90 (23): 11212–16. <https://doi.org/10.1073/pnas.90.23.11212>.

- Cassan-Wang, Hua, Nadia Goué, Mohammed N. Saidi, Sylvain Legay, Pierre Sivadon, Deborah Goffner, and Jacqueline Grima-Pettenati. 2013. "Identification of Novel Transcription Factors Regulating Secondary Cell Wall Formation in Arabidopsis." *Frontiers in Plant Science* 4. <https://doi.org/10.3389/fpls.2013.00189>.
- Casto, Anna L., Brian A. McKinley, Ka Man Jasmine Yu, William L. Rooney, and John E. Mullet. 2018. "Sorghum Stem Aerenchyma Formation Is Regulated by SbNAC_D during Internode Development." *Plant Direct* 2 (11): e00085. <https://doi.org/10.1002/pld3.85>.
- Chan, Ming-Tsair, Hsin-Hsiung Chang, Shin-Lon Ho, Wu-Fu Tong, and Su-May Yu. 1993. "Agrobacterium-Mediated Production of Transgenic Rice Plants Expressing a Chimeric β -Amylase Promoter/ β -Glucuronidase Gene." *Plant Molecular Biology* 22 (3): 491–506. <https://doi.org/10.1007/BF00015978>.
- Chantereau, J., Kondombo, C. "Estimation du taux d'allogamie chez les sorghos de la race guinea." In: Progress food grain research and production in semi-arid Africa, SAFGRAD: Ouagadougou, 1994; pp. 309-313.
- Chantereau, J., Cruz, J. F., Ratnadass, A., et Trouche G. (avec la collaboration de Flidel, G.), "Le sorgho." Versailles-Gembloux, Éditions Quæ-Presses agronomiques de Gembloux, « Agricultures tropicales en poche », 2013, 245 p.
- Char, Si Nian, Jialu Wei, Qi Mu, Xianran Li, Zhanyuan J. Zhang, Jianming Yu, and Bing Yang. 2019. "An Agrobacterium -delivered CRISPR /Cas9 System for Targeted Mutagenesis in Sorghum ." *Plant Biotechnology Journal*, 1–3. <https://doi.org/10.1111/pbi.13229>.
- Chauhan, Harsh, and Paramjit Khurana. 2017. "Wheat Genetic Transformation Using Mature Embryos as Explants." In *Methods in Molecular Biology*, 153–67. https://doi.org/10.1007/978-1-4939-7337-8_10.
- Chaw, Shu-Miaw, Chien-Chang Chang, Hsin-Liang Chen, and Wen-Hsiung Li. 2004. "Dating the Monocot/Dicot Divergence and the Origin of Core Eudicots Using Whole Chloroplast Genomes." *Journal of Molecular Evolution* 58 (4): 424–41. <https://doi.org/10.1007/s00239-003-2564-9>.
- Che, Ping, Ajith Anand, Emily Wu, Jeffry D. Sander, Marissa K. Simon, Weiwei Zhu, Amy L. Sigmund, et al. 2018. "Developing a Flexible, High-Efficiency Agrobacterium-Mediated Sorghum Transformation System with Broad Application." *Plant Biotechnology Journal* 16 (7): 1388–95. <https://doi.org/10.1111/pbi.12879>.
- Chen, Jiang, Qiang Yi, Qiaoheng Song, Yong Gu, Junjie Zhang, Yufeng Hu, Hanmei Liu, Yinghong Liu, Guowu Yu, and Yubi Huang. 2015. "A Highly Efficient Maize Nucellus Protoplast System for Transient Gene Expression and Studying Programmed Cell Death-Related Processes." *Plant Cell Reports* 34 (7): 1239–51. <https://doi.org/10.1007/s00299-015-1783-z>.
- Chen, Kunling, Yanpeng Wang, Rui Zhang, Huawei Zhang, and Caixia Gao. 2019. "CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture." *Annual Review of Plant Biology*

- 70 (1): 667–97. <https://doi.org/10.1146/annurev-arplant-050718-100049>.
- Chen, L., S. Zhang, R. N. Beachy, and C. M. Fauquet. 1998. “A Protocol for Consistent, Large-Scale Production of Fertile Transgenic Rice Plants.” *Plant Cell Reports* 18 (1–2): 25–31. <https://doi.org/10.1007/s002990050526>.
- Chen, Songbiao, Lizen Tao, Lirong Zeng, Miguel E. Vega-Sanchez, Kenji Umemura, and Guo Liang Wang. 2006. “A Highly Efficient Transient Protoplast System for Analyzing Defence Gene Expression and Protein-Protein Interactions in Rice.” *Molecular Plant Pathology* 7 (5): 417–27. <https://doi.org/10.1111/j.1364-3703.2006.00346.x>.
- Chiniquy, D., V. Sharma, A. Schultink, E. E. Baidoo, C. Rautengarten, K. Cheng, A. Carroll, et al. 2012. “XAX1 from Glycosyltransferase Family 61 Mediates Xylosyltransfer to Rice Xylan.” *Proceedings of the National Academy of Sciences* 109 (42): 17117–22. <https://doi.org/10.1073/pnas.1202079109>.
- Chinnusamy, Viswanathan, Jianhua Zhu, and Jian-Kang Zhu. 2007. “Cold Stress Regulation of Gene Expression in Plants.” *Trends in Plant Science* 12 (10): 444–51. <https://doi.org/10.1016/j.tplants.2007.07.002>.
- Cho, H T, and H Kende. 1997. “Expression of Expansin Genes Is Correlated with Growth in Deepwater Rice.” *The Plant Cell* 9 (9): 1661–71. <https://doi.org/10.1105/tpc.9.9.1661>.
- Chomzynski, P, and N Sacchi. 1987. “Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate–Phenol–Chloroform Extraction.” *Analytical Biochemistry* 162 (1): 156–59. <https://doi.org/10.1006/abio.1987.9999>.
- Christensen, A H, and P H Quail. 1996. “Ubiquitin Promoter-Based Vectors for High-Level Expression of Selectable and/or Screenable Marker Genes in Monocotyledonous Plants.” *Transgenic Research* 5 (3): 213–18.
- Christou, Paul. 1992. “Genetic Transformation of Crop Plants Using Microprojectile Bombardment.” *The Plant Journal* 2 (3): 275–81. <https://doi.org/10.1111/j.1365-313X.1992.00275.x>.
- Cleene, Marcel De, and Jozef De Ley. 1976. “The Host Range of Crown Gall.” *The Botanical Review* 42 (4): 389–466. <https://doi.org/10.1007/BF02860827>.
- Clerget, B, M Dingkuhn, J Chanterreau, J Hemberger, G Louarn, and M Vaksman. 2004. “Does Panicle Initiation in Tropical Sorghum Depend on Day-to-Day Change in Photoperiod?” *Field Crops Research* 88 (1): 21–37. <https://doi.org/10.1016/j.fcr.2003.11.008>.
- Clough, Steven J., and Andrew F. Bent. 1998. “Floral Dip: A Simplified Method for Agrobacterium-Mediated Transformation of Arabidopsis Thaliana.” *Plant Journal* 16 (6): 735–43. <https://doi.org/10.1046/j.1365-313X.1998.00343.x>.
- Cocking, Edward C. 1985. “Protoplasts from Root Hairs of Crop Plants.” *Nature Biotechnology* 3 (12): 1104–6. <https://doi.org/10.1038/nbt1285-1104>.
- Cone, Karen C., Suzy M. Cocciolone, Frances A. Burr, and Benjamin Burr. 1993. “Maize Anthocyanin Regulatory Gene P1 Is a Duplicate of C1 That Functions in the Plant.” *The Plant*

- Cell* 5 (12): 1795. <https://doi.org/10.2307/3869695>.
- Corcoran, Michael P., Diane L. McKay, and Jeffrey B. Blumberg. 2012. "Flavonoid Basics: Chemistry, Sources, Mechanisms of Action, and Safety." *Journal of Nutrition in Gerontology and Geriatrics* 31 (3): 176–89. <https://doi.org/10.1080/21551197.2012.698219>.
- Cornejo, Maria-Jesus, Diane Luth, Kathleen M. Blankenship, Olin D. Anderson, and Ann E. Blechl. 1993. "Activity of a Maize Ubiquitin Promoter in Transgenic Rice." *Plant Molecular Biology* 23 (3): 567–81. <https://doi.org/10.1007/BF00019304>.
- Cosgrove, Daniel J. 1997. "Assembly and enlargement of the primary cell wall in plants." *Annual Review of Cell and Developmental Biology* 13 (1): 171–201. <https://doi.org/10.1146/annurev.cellbio.13.1.171>.
- Cosgrove, Daniel J. 2000. "Loosening of Plant Cell Walls by Expansins." *Nature* 407 (6802): 321–26. <https://doi.org/10.1038/35030000>.
- Cosgrove, Daniel J. 2016. "Plant Cell Wall Extensibility: Connecting Plant Cell Growth with Cell Wall Structure, Mechanics, and the Action of Wall-Modifying Enzymes." *Journal of Experimental Botany* 67 (2): 463–76. <https://doi.org/10.1093/jxb/erv511>.
- Cosgrove, Daniel J. 2018. "Diffuse Growth of Plant Cell Walls." *Plant Physiology* 176 (1): 16–27. <https://doi.org/10.1104/pp.17.01541>.
- Daayf, Fouad, Abdelbasset El Hadrami, Ahmed F. El-Bebany, Maria A. Henriquez, Zhen Yao, Holly Derksen, Ismaïl El-Hadrami, and Lorne R. Adam. 2012. "Phenolic Compounds in Plant Defense and Pathogen Counter-Defense Mechanisms." In *Recent Advances in Polyphenol Research*, 191–208. Oxford, UK: Wiley-Blackwell. <https://doi.org/10.1002/9781118299753.ch8>.
- Dai, Shunhong, Ping Zheng, Philippe Marmey, Shiping Zhang, Wenzhong Tian, Shouyi Chen, Roger N. Beachy, and Claude Fauquet. 2001. "Comparative Analysis of Transgenic Rice Plants Obtained by Agrobacterium-Mediated Transformation and Particle Bombardment." *Molecular Breeding* 7 (1): 25–33. <https://doi.org/10.1023/A:1009687511633>.
- Dalton, S.J. 1988. "Plant Regeneration from Cell Suspension Protoplasts of *Festuca Arundinacea* Schreb. (Tall Fescue) and *Lolium Perenne* L. (Perennial Ryegrass)." *Journal of Plant Physiology* 132 (2): 170–75. [https://doi.org/10.1016/S0176-1617\(88\)80156-1](https://doi.org/10.1016/S0176-1617(88)80156-1).
- Dam, Sipko van, Urmo Vösa, Adriaan van der Graaf, Lude Franke, and João Pedro de Magalhães. 2018. "Gene Co-Expression Analysis for Functional Classification and Gene-Disease Predictions." *Briefings in Bioinformatics* 19 (4): 575–92. <https://doi.org/10.1093/bib/bbw139>.
- Davey, Michael R., Paul Anthony, J. Brian Power, and Kenneth C. Lowe. 2005. "Plant Protoplasts: Status and Biotechnological Perspectives." *Biotechnology Advances* 23 (2): 131–71. <https://doi.org/10.1016/j.biotechadv.2004.09.008>.
- Davidson, Rebecca M., Malali Gowda, Gaurav Moghe, Haining Lin, Brieanne Vaillancourt, Shin-Han Shiu, Ning Jiang, and C. Robin Buell. 2012. "Comparative Transcriptomics of Three Poaceae Species Reveals Patterns of Gene Expression Evolution." *The Plant Journal*, June, no-no.

<https://doi.org/10.1111/j.1365-313X.2012.05005.x>.

- Davin, Laurence B., and Norman G. Lewis. 2000. "Dirigent Proteins and Dirigent Sites Explain the Mystery of Specificity of Radical Precursor Coupling in Lignan and Lignin Biosynthesis." *Plant Physiology* 123 (2): 453–62. <https://doi.org/10.1104/pp.123.2.453>.
- Delalande M., Regnard J. L., Costes E., Reymond M., Luquet D., Fabre D., et al. (2015). DiaPHEN: Field Platform to Analyze Mechanisms of Drought Tolerance at Field Level. Montpellier: EUCARPIA.
- Delmer, D P, and Y Amor. 1995. "Cellulose Biosynthesis." *The Plant Cell* 7 (7): 987–1000. <https://doi.org/10.1105/tpc.7.7.987>.
- Delmer, Deborah P., and Candace H. Haigler. 2002. "The Regulation of Metabolic Flux to Cellulose, a Major Sink for Carbon in Plants." *Metabolic Engineering* 4 (1): 22–28. <https://doi.org/10.1006/mben.2001.0206>.
- Demura, Taku, and Zheng-Hua Ye. 2010. "Regulation of Plant Biomass Production." *Current Opinion in Plant Biology* 13 (3): 298–303. <https://doi.org/10.1016/j.pbi.2010.03.002>.
- DeRose, Richard T., Dilara Begum, and Timothy C. Hall. 1996. "Analysis of Kafirin Promoter Activity in Transgenic Tobacco Seeds." *Plant Molecular Biology* 32 (6): 1029–35. <https://doi.org/10.1007/BF00041386>.
- Deu, M, F Rattunde, and J Chantreau. 2006. "A Global View of Genetic Diversity in Cultivated Sorghums Using a Core Collection." *Genome* 49 (2): 168–80. <https://doi.org/10.1139/g05-092>.
- Dharanishanthi, Veeramuthu, and Modhumita Ghosh Dasgupta. 2018. "Co-Expression Network of Transcription Factors Reveal Ethylene-Responsive Element-Binding Factor as Key Regulator of Wood Phenotype in Eucalyptus Tereticornis." *3 Biotech* 8 (7): 315. <https://doi.org/10.1007/s13205-018-1344-6>.
- Dhugga, K. S. 2004. "Guar Seed -Mannan Synthase Is a Member of the Cellulose Synthase Super Gene Family." *Science* 303 (5656): 363–66. <https://doi.org/10.1126/science.1090908>.
- Dien, Bruce S., Gautam Sarath, Jeffrey F. Pedersen, Scott E. Sattler, Han Chen, Deanna L. Funnell-Harris, Nancy N. Nichols, and Michael A. Cotta. 2009. "Improved Sugar Conversion and Ethanol Yield for Forage Sorghum (*Sorghum Bicolor* L. Moench) Lines with Reduced Lignin Contents." *Bioenergy Research* 2 (3): 153–64. <https://doi.org/10.1007/s12155-009-9041-2>.
- Ding, Shi-You, and Michael E. Himmel. 2006. "The Maize Primary Cell Wall Microfibril: A New Model Derived from Direct Visualization." *Journal of Agricultural and Food Chemistry* 54 (3): 597–606. <https://doi.org/10.1021/jf051851z>.
- Ding, Zehong, Sarit Weissmann, Minghui Wang, Baijuan Du, Lei Huang, Lin Wang, Xiaoyu Tu, et al. 2015. "Identification of Photosynthesis-Associated C4 Candidate Genes through Comparative Leaf Gradient Transcriptome in Multiple Lineages of C3 and C4 Species." *PLoS ONE* 10 (10): 1–19. <https://doi.org/10.1371/journal.pone.0140629>.
- Dixon, R. A., and N. L. Paiva. 1995. "Stress-Induced Phenylpropanoid Metabolism." *The Plant Cell*,

- July, 1085–97. <https://doi.org/10.1105/tpc.7.7.1085>.
- Dixon, Richard A., Lahoucine Achnine, Parvathi Kota, Chang-Jun Liu, M. S. Srinivasa Reddy, and Liangjiang Wang. 2002. “The Phenylpropanoid Pathway and Plant Defence-a Genomics Perspective.” *Molecular Plant Pathology* 3 (5): 371–90. <https://doi.org/10.1046/j.1364-3703.2002.00131.x>.
- Do, Phat T., Hyeyoung Lee, Muruganantham Mookkan, William R. Folk, and Zhanyuan J. Zhang. 2016. “Rapid and Efficient Agrobacterium-Mediated Transformation of Sorghum (*Sorghum Bicolor*) Employing Standard Binary Vectors and Bar Gene as a Selectable Marker.” *Plant Cell Reports* 35 (10): 2065–76. <https://doi.org/10.1007/s00299-016-2019-6>.
- Doblin, Monika S., Isaac Kurek, Deborah Jacob-Wilk, and Deborah P. Delmer. 2002. “Cellulose Biosynthesis in Plants: From Genes to Rosettes.” *Plant and Cell Physiology* 43 (12): 1407–20. <https://doi.org/10.1093/pcp/pcf164>.
- Doblin, Monika S., Filomena A. Pettolino, Sarah M. Wilson, Rebecca Campbell, Rachel A. Burton, Geoffrey B. Fincher, Ed Newbigin, and Antony Bacic. 2009. “A Barley Cellulose Synthase-like CSLH Gene Mediates (1,3;1,4)- β -D- Glucan Synthesis in Transgenic Arabidopsis.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (14): 5996–6001. <https://doi.org/10.1073/pnas.0902019106>.
- Doblin, Monika S., Filomena Pettolino, and Antony Bacic. 2010. “Evans Review: Plant Cell Walls: The Skeleton of the Plant World.” *Functional Plant Biology* 37 (5): 357–81. <https://doi.org/10.1071/FP09279>.
- Doench, John G, Nicolo Fusi, Meagan Sullender, Mudra Hegde, Emma W Vaimberg, Katherine F Donovan, Ian Smith, et al. 2016. “Optimized SgRNA Design to Maximize Activity and Minimize Off-Target Effects of CRISPR-Cas9.” *Nature Biotechnology* 34 (2): 184–91. <https://doi.org/10.1038/nbt.3437>.
- Dong, Jin-Zhuo, Mei-Zhu Yang, Shi-Rong Jia, and Nam-Hai Chua. 1991. “Transformation of Melon (*Cucumis Melo* L.) and Expression from the Cauliflower Mosaic Virus 35S Promoter in Transgenic Melon Plants.” *Nature Biotechnology* 9 (9): 858–63. <https://doi.org/10.1038/nbt0991-858>.
- Du, Hai, Bo-run Feng, Si-si Yang, Yu-bi Huang, and Yi-xiong Tang. 2012. “The R2R3-MYB Transcription Factor Gene Family in Maize” 7 (6): 1–12. <https://doi.org/10.1371/journal.pone.0037463>.
- Dubos, Christian, José Le Gourriec, Antoine Baudry, Gunnar Huep, Elodie Lanet, Isabelle Debeaujon, Jean-Marc Routaboul, Alessandro Alboresi, Bernd Weisshaar, and Loïc Lepiniec. 2008. “MYBL2 Is a New Regulator of Flavonoid Biosynthesis in Arabidopsis Thaliana.” *The Plant Journal* 55 (6): 940–53. <https://doi.org/10.1111/j.1365-313X.2008.03564.x>.
- Dubos, Christian, Ralf Stracke, Erich Grotewold, Bernd Weisshaar, Cathie Martin, and Loïc Lepiniec. 2010. “MYB Transcription Factors in Arabidopsis.” *Trends in Plant Science* 15 (10): 573–81.

- <https://doi.org/10.1016/j.tplants.2010.06.005>.
- Dufayard, J.-F., L. Duret, S. Penel, M. Gouy, F. Rechenmann, and G. Perriere. 2005. "Tree Pattern Matching in Phylogenetic Trees: Automatic Search for Orthologs or Paralogs in Homologous Gene Sequence Databases." *Bioinformatics* 21 (11): 2596–2603. <https://doi.org/10.1093/bioinformatics/bti325>.
- Dugas, Diana V, Marcela K Monaco, Andrew Olson, Robert R Klein, Sunita Kumari, Doreen Ware, and Patricia E Klein. 2011. "Functional Annotation of the Transcriptome of Sorghum Bicolor in Response to Osmotic Stress and Absciscic Acid." *BMC Genomics* 12 (1): 514. <https://doi.org/10.1186/1471-2164-12-514>.
- Dwivedi, Krishna K., Dominique Roche, and John G. Carman. 2010. "Expression in Arabidopsis of a Nucellus-Specific Promoter from Watermelon (*Citrullus Lanatus*)." *Plant Science* 179 (5): 549–52. <https://doi.org/10.1016/j.plantsci.2010.07.016>.
- Dykes, Linda, and Lloyd W. Rooney. 2006. "Sorghum and Millet Phenols and Antioxidants." *Journal of Cereal Science* 44 (3): 236–51. <https://doi.org/10.1016/j.jcs.2006.06.007>.
- Ebringerová, Anna. 2006. "Structural Diversity and Application Potential of Hemicelluloses." *Macromolecular Symposia* 232 (333): 1–12. <https://doi.org/10.1002/masy.200551401>.
- Ebringerová, Anna, Zdenka Hromádková, and Thomas Heinze. 2005. "Hemicellulose." In *Polysaccharides I*, 1–67. Berlin/Heidelberg: Springer-Verlag. <https://doi.org/10.1007/b136816>.
- Eckardt, Nancy A. 2003. "Cellulose Synthesis Takes the Cesa Train." *The Plant Cell* 15 (8): 1685–87. <https://doi.org/10.1105/tpc.150810>.
- Elkonin, Lev A., Julia V. Italyanskaya, Valery M. Panin, and Nikolai Yu. Selivanov. 2017. "Development of Transgenic Sorghum Plants with Improved In Vitro Kafirin Digestibility." In *Plant Engineering*. InTech. <https://doi.org/10.5772/intechopen.69973>.
- Elshire, Robert J., Jeffrey C. Glaubitz, Qi Sun, Jesse A. Poland, Ken Kawamoto, Edward S. Buckler, and Sharon E. Mitchell. 2011. "A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species." Edited by Laszlo Orban. *PLoS ONE* 6 (5): e19379. <https://doi.org/10.1371/journal.pone.0019379>.
- Enciso, J., J. Jifon, L. Ribera, S.D. Zapata, and G.K. Ganjegunte. 2015. "Yield, Water Use Efficiency and Economic Analysis of Energy Sorghum in South Texas." *Biomass and Bioenergy* 81 (October): 339–44. <https://doi.org/10.1016/j.biombioe.2015.07.021>.
- Endo, Hitoshi, Masatoshi Yamaguchi, Taizo Tamura, Yoshimi Nakano, Nobuyuki Nishikubo, Arata Yoneda, Ko Kato, et al. 2015. "Multiple Classes of Transcription Factors Regulate the Expression of VASCULAR-RELATED NAC-DOMAIN7, a Master Switch of Xylem Vessel Differentiation." *Plant and Cell Physiology* 56 (2): 242–54. <https://doi.org/10.1093/pcp/pcu134>.
- Ermawar, Riksfardini A., Helen M. Collins, Caitlin S. Byrt, Natalie S. Betts, Marilyn Henderson, Neil J. Shirley, Julian Schwerdt, Jelle Lahnstein, Geoffrey B. Fincher, and Rachel A. Burton. 2015. "Distribution, Structure and Biosynthetic Gene Families of (1,3;1,4)- β -Glucan in Sorghum

- Bicolor.” *Journal of Integrative Plant Biology* 57 (4): 429–45. <https://doi.org/10.1111/jipb.12338>.
- Ermawar, Riksfardini A, Helen M Collins, Caitlin S Byrt, Marilyn Henderson, Lisa A O’Donovan, Neil J Shirley, Julian G Schwerdt, Jelle Lahnstein, Geoffrey B Fincher, and Rachel A Burton. 2015. “Genetics and Physiology of Cell Wall Polysaccharides in the Model C4 Grass, *Setaria Viridis* Spp.” *BMC Plant Biology* 15: 236. <https://doi.org/10.1186/s12870-015-0624-0>.
- Ermeýdan, Mahmut A., Etienne Cabane, Admir Masic, Joachim Koetz, and Ingo Burgert. 2012. “Flavonoid Insertion into Cell Walls Improves Wood Properties.” *ACS Applied Materials & Interfaces* 4 (11): 5782–89. <https://doi.org/10.1021/am301266k>.
- Esquerré-Tugayé, Marie-Thérèse, Georges Boudart, and Bernard Dumas. 2000. “Cell Wall Degrading Enzymes, Inhibitory Proteins, and Oligosaccharides Participate in the Molecular Dialogue between Plants and Pathogens.” *Plant Physiology and Biochemistry* 38 (1–2): 157–63. [https://doi.org/10.1016/S0981-9428\(00\)00161-3](https://doi.org/10.1016/S0981-9428(00)00161-3).
- Fan, Kai, Ming Wang, Ying Miao, Mi Ni, Noreen Bibi, Shuna Yuan, Feng Li, and Xuede Wang. 2014. “Molecular Evolution and Expansion Analysis of the NAC Transcription Factor in *Zea Mays*” 9 (11). <https://doi.org/10.1371/journal.pone.0111837>.
- Fan, Ling, Raphael Linker, Shimon Gepstein, Eiichi Tanimoto, Ryoichi Yamamoto, and Peter M. Neumann. 2006. “Progressive Inhibition by Water Deficit of Cell Wall Extensibility and Growth along the Elongation Zone of Maize Roots Is Related to Increased Lignin Metabolism and Progressive Stelar Accumulation of Wall Phenolics.” *Plant Physiology* 140 (2): 603–12. <https://doi.org/10.1104/pp.105.073130>.
- Fang, Yujie, Jun You, Kabin Xie, Weibo Xie, and Lizhong Xiong. 2008. “Systematic Sequence Analysis and Identification of Tissue-Specific or Stress-Responsive Genes of NAC Transcription Factor Family in Rice.” *Molecular Genetics and Genomics* 280 (6): 547–63. <https://doi.org/10.1007/s00438-008-0386-6>.
- Farré, Imma, and José María Faci. 2006. “Comparative Response of Maize (*Zea Mays* L.) and Sorghum (*Sorghum Bicolor* L. Moench) to Deficit Irrigation in a Mediterranean Environment.” *Agricultural Water Management* 83 (1–2): 135–43. <https://doi.org/10.1016/j.agwat.2005.11.001>.
- Festucci-Buselli, Reginaldo A., Wagner C. Otoni, and Chandrashekhar P. Joshi. 2007. “Structure, Organization, and Functions of Cellulose Synthase Complexes in Higher Plants.” *Brazilian Journal of Plant Physiology* 19 (1): 1–13. <https://doi.org/10.1590/s1677-04202007000100001>.
- Ficklin, Stephen P., and F. Alex Feltus. 2011. “Gene Coexpression Network Alignment and Conservation of Gene Modules between Two Grass Species: Maize and Rice.” *Plant Physiology* 156 (3): 1244–56. <https://doi.org/10.1104/pp.111.173047>.
- Fornalé, Silvia, Eric Lopez, Jorge E. Salazar-Henao, Pedro Fernández-Nohales, Joan Rigau, and David Caparros-Ruiz. 2014. “AtMYB7, a New Player in the Regulation of UV-Sunscreens in *Arabidopsis Thaliana*.” *Plant and Cell Physiology* 55 (3): 507–16.

<https://doi.org/10.1093/pcp/pct187>.

- Fornalé, Silvia, Xinhui Shi, Chenglin Chai, Antonio Encina, Sami Irar, Montserrat Capellades, Elisabet Fuguet, et al. 2010. "ZmMYB31 Directly Represses Maize Lignin Genes and Redirects the Phenylpropanoid Metabolic Flux." *The Plant Journal* 64 (4): 633–44. <https://doi.org/10.1111/j.1365-313X.2010.04363.x>.
- Fornalé, Silvia, Fathi-Mohamed Sonbol, Tamara Maes, Montserrat Capellades, Pere Puigdomènech, Joan Rigau, and David Caparrós-Ruiz. 2006. "Down-Regulation of the Maize and Arabidopsis Thaliana Caffeic Acid O-Methyl-Transferase Genes by Two New Maize R2R3-MYB Transcription Factors." *Plant Molecular Biology* 62 (6): 809–23. <https://doi.org/10.1007/s11103-006-9058-2>.
- Fraley, R. T., S. G. Rogers, R. B. Horsch, P. R. Sanders, J. S. Flick, S. P. Adams, M. L. Bittner, et al. 1983. "Expression of Bacterial Genes in Plant Cells." *Proceedings of the National Academy of Sciences* 80 (15): 4803–7. <https://doi.org/10.1073/pnas.80.15.4803>.
- Fraser, Christopher M., and Clint Chapple. 2011. "The Phenylpropanoid Pathway in Arabidopsis." *The Arabidopsis Book* 9 (January): e0152. <https://doi.org/10.1199/tab.0152>.
- Frère, Céline H., Peter J. Prentis, Edward K. Gilding, Agnieszka M. Mudge, Alan Cruickshank, and Ian D. Godwin. 2011. "Lack of Low Frequency Variants Masks Patterns of Non-Neutral Evolution Following Domestication." Edited by Gro V. Amdam. *PLoS ONE* 6 (8): e23041. <https://doi.org/10.1371/journal.pone.0023041>.
- Fromm, Michael E., Fionnuala Morrish, Charles Armstrong, Rosalind Williams, John Thomas, and Theodore M. Klein. 1990. "Inheritance and Expression of Chimeric Genes in the Progeny of Transgenic Maize Plants." *Nature Biotechnology* 8 (9): 833–39. <https://doi.org/10.1038/nbt0990-833>.
- Fromm, Michael E., Loverine P. Taylor, and Virginia Walbot. 1986. "Stable Transformation of Maize after Gene Transfer by Electroporation." *Nature* 319 (6056): 791–93. <https://doi.org/10.1038/319791a0>.
- Fry, Stephen C. 1998. "Oxidative Scission of Plant Cell Wall Polysaccharides by Ascorbate-Induced Hydroxyl Radicals." *Biochemical Journal* 332 (2): 507–15. <https://doi.org/10.1042/bj3320507>.
- Fujikawa, Yukichi, Takahiro Nakanishi, Hiroko Kawakami, Kanako Yamasaki, Masa H Sato, Hiroyuki Tsuji, Makoto Matsuoka, and Naohiro Kato. 2014. "Split Luciferase Complementation Assay to Detect Regulated Protein-Protein Interactions in Rice Protoplasts in a Large-Scale Format." *Rice* 7 (1): 11. <https://doi.org/10.1186/s12284-014-0011-8>.
- Fujita, Miki, Yasunari Fujita, Kyonoshin Maruyama, Motoaki Seki, Keiichiro Hiratsu, Masaru Ohme-Takagi, Lam-Son Phan Tran, Kazuko Yamaguchi-Shinozaki, and Kazuo Shinozaki. 2004. "A Dehydration-Induced NAC Protein, RD26, Is Involved in a Novel ABA-Dependent Stress-Signaling Pathway." *The Plant Journal* 39 (6): 863–76. <https://doi.org/10.1111/j.1365-313X.2004.02171.x>.

- Fujiwara, Sumire, and Nobutaka Mitsuda. 2016. "ANAC075,a Putative Regulator of VASCULAR-RELATED NAC-DOMAIN7,Is a Repressor of Flowering." *Plant Biotechnology* 33 (4): 255–65. <https://doi.org/10.5511/plantbiotechnology.16.0215b>.
- Fukao, Takeshi, Kenong Xu, Pamela C. Ronald, and Julia Bailey-Serres. 2006. "A Variable Cluster of Ethylene Response Factor–Like Genes Regulates Metabolic and Developmental Acclimation Responses to Submergence in Rice." *The Plant Cell* 18 (8): 2021–34. <https://doi.org/10.1105/tpc.106.043000>.
- Furtwengler, Pierre, and Luc Avérous. 2018. "From D-Sorbitol to Five-Membered Bis(Cyclo-Carbonate) as a Platform Molecule for the Synthesis of Different Original Biobased Chemicals and Polymers." *Scientific Reports* 8 (1): 9134. <https://doi.org/10.1038/s41598-018-27450-w>.
- Gallego-Giraldo, Lina, Gail Shadle, Hui Shen, Jaime Barros-Rios, Sandra Fresquet Corrales, Huanzhong Wang, and Richard A. Dixon. 2016. "Combining Enhanced Biomass Density with Reduced Lignin Level for Improved Forage Quality." *Plant Biotechnology Journal* 14 (3): 895–904. <https://doi.org/10.1111/pbi.12439>.
- Gamazon, Eric R., Heather E. Wheeler, Kaanan P. Shah, Sahar V. Mozaffari, Keston Aquino-Michaels, Robert J. Carroll, Anne E. Eyler, et al. 2015. "A Gene-Based Association Method for Mapping Traits Using Reference Transcriptome Data." *Nature Genetics* 47 (9): 1091–98. <https://doi.org/10.1038/ng.3367>.
- Gao, Zhensheng, Xueju Xie, Yan Ling, Subbarat Muthukrishnan, and George H. Liang. 2005. "Agrobacterium Tumefaciens-Mediated Sorghum Transformation Using a Mannose Selection System." *Plant Biotechnology Journal* 3 (6): 591–99. <https://doi.org/10.1111/j.1467-7652.2005.00150.x>.
- Gardner, J. C., J. W. Maranville, and E. T. Paparozzi. 1994. "Nitrogen Use Efficiency among Diverse Sorghum Cultivars." *Crop Science* 34 (3): 728. <https://doi.org/10.2135/cropsci1994.0011183X003400030023x>.
- Garofalo, P., and M. Rinaldi. 2013. "Water-Use Efficiency of Irrigated Biomass Sorghum in a Mediterranean Environment." *Spanish Journal of Agricultural Research* 11 (4): 1153. <https://doi.org/10.5424/sjar/2013114-4147>.
- Garsmeur, Olivier, Gaetan Droc, Rudie Antonise, Jane Grimwood, Bernard Potier, Karen Aitken, Jerry Jenkins, et al. 2018. "A Mosaic Monoploid Reference Sequence for the Highly Complex Genome of Sugarcane." *Nature Communications* 9 (1): 2638. <https://doi.org/10.1038/s41467-018-05051-5>.
- Gasser, C. S., and R. T. Fraley. 1989. "Genetically Engineering Plants for Crop Improvement." *Science* 244 (4910): 1293–99. <https://doi.org/10.1126/science.244.4910.1293>.
- Gatz, Christine. 1996. "Chemically Inducible Promoters in Transgenic Plants." *Current Opinion in Biotechnology* 7 (2): 168–72. [https://doi.org/10.1016/S0958-1669\(96\)80008-5](https://doi.org/10.1016/S0958-1669(96)80008-5).
- Gelli, Malleswari, Yongchao Duo, Anji Konda, Chi Zhang, David Holding, and Ismail Dweikat. 2014.

- “Identification of Differentially Expressed Genes between Sorghum Genotypes with Contrasting Nitrogen Stress Tolerance by Genome-Wide Transcriptional Profiling.” *BMC Genomics* 15 (1): 179. <https://doi.org/10.1186/1471-2164-15-179>.
- Gelvin, Stanton B. 2003. “Plant Transformation: The Biology behind the ‘Gene-Jockeying’ Tool.” *Society* 67 (1): 16–37. <https://doi.org/10.1128/MMBR.67.1.16>.
- Gille, Sascha, and Markus Pauly. 2012. “O-Acetylation of Plant Cell Wall Polysaccharides.” *Frontiers in Plant Science* 3 (JAN): 1–7. <https://doi.org/10.3389/fpls.2012.00012>.
- Girijashankar, V., H. C. Sharma, Kiran K. Sharma, V. Swathisree, L. Sivarama Prasad, B. V. Bhat, Monique Royer, et al. 2005. “Development of Transgenic Sorghum for Insect Resistance against the Spotted Stem Borer (*Chilo Partellus*).” *Plant Cell Reports* 24 (9): 513–22. <https://doi.org/10.1007/s00299-005-0947-7>.
- Girijashankar, V., and V. Swathisree. 2009. “Genetic Transformation of Sorghum Bicolor.” *Physiology and Molecular Biology of Plants* 15 (4): 287–302. <https://doi.org/10.1007/s12298-009-0033-7>.
- Goicoechea, Monica, Eric Lacombe, Sylvain Legay, Snjezana Mihaljevic, Philippe Rech, Alain Jauneau, Catherine Lapierre, et al. 2005. “EgMYB2, a New Transcriptional Activator from Eucalyptus Xylem, Regulates Secondary Cell Wall Formation and Lignin Biosynthesis.” *Plant Journal* 43 (4): 553–67. <https://doi.org/10.1111/j.1365-313X.2005.02480.x>.
- Golovkin, Maxim V., Mariann Ábrahám, Sándor Mórocz, Sándor Bottka, Attila Fehér, and Dénes Dudits. 1993. “Production of Transgenic Maize Plants by Direct DNA Uptake into Embryogenic Protoplasts.” *Plant Science* 90 (1): 41–52. [https://doi.org/10.1016/0168-9452\(93\)90154-R](https://doi.org/10.1016/0168-9452(93)90154-R).
- Guan, Qingmei, Xiule Yue, Haitao Zeng, and Jianhua Zhu. 2014. “The Protein Phosphatase RCF2 and Its Interacting Partner NAC019 Are Critical for Heat Stress–Responsive Gene Regulation and Thermotolerance in *Arabidopsis*.” *The Plant Cell* 26 (1): 438–53. <https://doi.org/10.1105/tpc.113.118927>.
- Guindon, Stéphane, Jean François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, and Olivier Gascuel. 2010. “New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0.” *Systematic Biology* 59 (3): 307–21. <https://doi.org/10.1093/sysbio/syq010>.
- Gupta, Chirag, and Andy Pereira. 2019. “Recent Advances in Gene Function Prediction Using Context-Specific Coexpression Networks in Plants [Version 1; Referees: 2 Approved].” *F1000Research* 8: 1–10. <https://doi.org/10.12688/f1000research.17207.1>.
- Gurel, Songul, Ekrem Gurel, Rajvinder Kaur, Joshua Wong, Ling Meng, Han Qi Tan, and Peggy G. Lemaux. 2009. “Efficient, Reproducible Agrobacterium-Mediated Transformation of Sorghum Using Heat Treatment of Immature Embryos.” *Plant Cell Reports* 28 (3): 429–44. <https://doi.org/10.1007/s00299-008-0655-1>.
- Gurel, Songul, Ekrem Gurel, Tamara I. Miller, and Peggy G. Lemaux. 2012. “Agrobacterium-

- Mediated Transformation of Sorghum Bicolor Using Immature Embryos.” In , 109–22. https://doi.org/10.1007/978-1-61779-558-9_10.
- Gutjahr, Sylvain, Michel Vaksman, Michaël Dingkuhn, Korothimi Thera, Gilles Trouche, Serge Braconnier, and Delphine Luquet. 2013. “Grain, Sugar and Biomass Accumulation in Tropical Sorghums. I. Trade-Offs and Effects of Phenological Plasticity.” *Functional Plant Biology* 40 (4): 342–54. <https://doi.org/10.1071/FP12269>.
- Haga, N., K. Kato, M. Murase, S. Araki, M. Kubo, T. Demura, K. Suzuki, et al. 2007. “R1R2R3-Myb Proteins Positively Regulate Cytokinesis through Activation of KNOLLE Transcription in Arabidopsis Thaliana.” *Development* 134 (6): 1101–10. <https://doi.org/10.1242/dev.02801>.
- Hagio, Takashi, Alan D. Blowers, and Elizabeth D. Earle. 1991. “Stable Transformation of Sorghum Cell Cultures after Bombardment with DNA-Coated Microprojectiles.” *Plant Cell Reports* 10 (5): 260–64. <https://doi.org/10.1007/BF00232571>.
- Haigler, Candace H., V. Michelle Babb, Sangjoon Hwang, and Vadim V. Salnikov. 2001. “Regulation of Cellulose Biosynthesis in Developing Xylem.” In , 1–9. [https://doi.org/10.1016/S0921-0423\(01\)80050-2](https://doi.org/10.1016/S0921-0423(01)80050-2).
- Hallam, A., I.C. Anderson, and D.R. Buxton. 2001. “Comparative Economic Analysis of Perennial, Annual, and Intercrops for Biomass Production.” *Biomass and Bioenergy* 21 (6): 407–24. [https://doi.org/10.1016/S0961-9534\(01\)00051-4](https://doi.org/10.1016/S0961-9534(01)00051-4).
- Hamada, Haruyasu, Qianyan Linghu, Yozo Nagira, Ryuji Miki, Naoaki Taoka, and Ryoza Imai. 2017. “An in Planta Biolistic Method for Stable Wheat Transformation.” *Scientific Reports* 7 (1): 2–9. <https://doi.org/10.1038/s41598-017-11936-0>.
- Hamid, Rifat, Saleem Javed, Mahboob Ahmad, Malik Mobeen Ahmad, Malik Zainul Abidin, Minhaj A Khan, and Javed Musarrat. 2013. “Chitinases: An Update.” *Journal of Pharmacy and Bioallied Sciences* 5 (1): 21. <https://doi.org/10.4103/0975-7406.106559>.
- Handakumbura, Pubudu P., and Samuel P. Hazen. 2012. “Transcriptional Regulation of Grass Secondary Cell Wall Biosynthesis: Playing Catch-Up with Arabidopsis Thaliana.” *Frontiers in Plant Science* 3. <https://doi.org/10.3389/fpls.2012.00074>.
- Handakumbura, Pubudu P, Dominick A Matos, Karen S Osmont, Michael J Harrington, Kyuyoung Heo, Kabindra Kafle, Seong H Kim, Tobias I Baskin, and Samuel P Hazen. 2013. “Perturbation of Brachypodium Distachyon CELLULOSE SYNTHASE A4 or 7 Results in Abnormal Cell Walls.” *BMC Plant Biology* 13 (1): 131. <https://doi.org/10.1186/1471-2229-13-131>.
- Hansen, Geneviève. 2000. “Evidence for Agrobacterium-Induced Apoptosis in Maize Cells.” *Molecular Plant-Microbe Interactions* 13 (6): 649–57. <https://doi.org/10.1094/MPMI.2000.13.6.649>.
- Hansen, and Wright. 1999. “Recent Advances in the Transformation of Plants.” *Trends in Plant Science* 4 (6): 226–31.
- Harris, Philip J., and Jason A.K. Trethewey. 2010. “The Distribution of Ester-Linked Ferulic Acid in

- the Cell Walls of Angiosperms.” *Phytochemistry Reviews* 9 (1): 19–33. <https://doi.org/10.1007/s11101-009-9146-4>.
- Harris, Rita, Martha Wright, Michael Byrne, James Varnum, Blanche Brightwell, and Karel Schubert. 1988. “Callus Formation and Plantlet Regeneration from Protoplasts Derived from Suspension Cultures of Wheat (*Triticum Aestivum* L.).” *Plant Cell Reports* 7 (5): 337–40. <https://doi.org/10.1007/BF00269932>.
- Harshavardhan, Doddapaneni, Thammiraju Shyamala RANI, Kandasamy Ulaganathan, and Nadoor Seetharama. 2002. “An Improved Protocol for Regeneration of Sorghum Bicolor from Isolated Shoot Apices.” *Plant Biotechnology* 19 (3): 163–71. <https://doi.org/10.5511/plantbiotechnology.19.163>.
- He, Xin-Jian, Rui-Ling Mu, Wan-Hong Cao, Zhi-Gang Zhang, Jin-Song Zhang, and Shou-Yi Chen. 2005. “AtNAC2, a Transcription Factor Downstream of Ethylene and Auxin Signaling Pathways, Is Involved in Salt Stress Response and Lateral Root Development.” *The Plant Journal* 44 (6): 903–16. <https://doi.org/10.1111/j.1365-313X.2005.02575.x>.
- Hennet, Lauriane, Angélique Berger, Noemi Trabanco, Emeline Ricciuti, Jean-François Dufayard, Stéphanie Bocs, Denis Bastianelli, et al. 2020. “Transcriptional Regulation of Sorghum Stem Composition: Key Players Identified Through Co-Expression Gene Network and Comparative Genomics Analyses.” *Frontiers in Plant Science* 11 (March). <https://doi.org/10.3389/fpls.2020.00224>.
- Hernández-Blanco, Camilo, Dong Xin Feng, Jian Hu, Andrea Sánchez-Vallet, Laurent Deslandes, Francisco Llorente, Marta Berrocal-Lobo, et al. 2007. “Impairment of Cellulose Synthases Required for Arabidopsis Secondary Cell Wall Formation Enhances Disease Resistance.” *Plant Cell* 19 (3): 890–903. <https://doi.org/10.1105/tpc.106.048058>.
- Hibara, Ken-ichiro, Shinobu Takada, and Masao Tasaka. 2003. “CUC1 Gene Activates the Expression of SAM-Related Genes to Induce Adventitious Shoot Formation.” *The Plant Journal* 36 (5): 687–96. <https://doi.org/10.1046/j.1365-313X.2003.01911.x>.
- Hibara, Ken Ichiro, Md Resaul Karim, Shinobu Takada, Ken Ichiro Taoka, Masahiko Furutani, Mitsuhiro Aida, and Masao Tasaka. 2006. “Arabidopsis CUP-SHAPED COTYLEDON3 Regulates Postembryonic Shoot Meristem and Organ Boundary Formation.” *Plant Cell* 18 (11): 2946–57. <https://doi.org/10.1105/tpc.106.045716>.
- Hiei, Yukoh, and Toshihiko Komari. 2008. “Agrobacterium-Mediated Transformation of Rice Using Immature Embryos or Calli Induced from Mature Seed.” *Nature Protocols* 3 (5): 824–34. <https://doi.org/10.1038/nprot.2008.46>.
- Hiei, Yukoh, Shozo Ohta, Toshihiko Komari, and Takashi Kumashiro. 1994. “Efficient Transformation of Rice (*Oryza Sativa* L.) Mediated by Agrobacterium and Sequence Analysis of the Boundaries of the T-DNA.” *The Plant Journal* 6 (2): 271–82. <https://doi.org/10.1046/j.1365-313X.1994.6020271.x>.

- Higashi, Yasuhiro, and Kazuki Saito. 2013. "Network Analysis for Gene Discovery in Plant-Specialized Metabolism." *Plant, Cell and Environment* 36 (9): 1597–1606. <https://doi.org/10.1111/pce.12069>.
- Hirano, Ko, Koichiro Aya, Yoichi Morinaka, Shiro Nagamatsu, Yutaka Sato, Baltazar A. Antonio, Nobukazu Namiki, Yoshiaki Nagamura, and Makoto Matsuoka. 2013a. "Survey of Genes Involved in Rice Secondary Cell Wall Formation Through a Co-Expression Network." *Plant and Cell Physiology* 54 (11): 1803–21. <https://doi.org/10.1093/pcp/pct121>.
- Hirano, Ko, Mari Kondo, Koichiro Aya, Akio Miyao, Yutaka Sato, Baltazar A. Antonio, Nobukazu Namiki, Yoshiaki Nagamura, and Makoto Matsuoka. 2013b. "Identification of Transcription Factors Involved in Rice Secondary Cell Wall Formation." *Plant and Cell Physiology* 54 (11): 1791–1802. <https://doi.org/10.1093/pcp/pct122>.
- Hoorn, Renier A. L. van der. 2008. "Plant Proteases: From Phenotypes to Molecular Mechanisms." *Annual Review of Plant Biology* 59 (1): 191–223. <https://doi.org/10.1146/annurev.arplant.59.032607.092835>.
- Horsch, Robert B, Robert T Fraley, Stephen G Rogers, Patricia R Sanders, and Nancy Hoffmann. 2016. "Inheritance of Functional Foreign Genes in Plants Published by : American Association for the Advancement of Science Stable URL : [Http://Www.Jstor.Org/Stable/1692099](http://www.jstor.org/stable/1692099)" 223 (4635): 496–98.
- Howe, Arlene, Shirley Sato, Ismail Dweikat, Mike Fromm, and Tom Clemente. 2006. "Rapid and Reproducible Agrobacterium-Mediated Transformation of Sorghum." *Plant Cell Reports* 25 (8): 784–91. <https://doi.org/10.1007/s00299-005-0081-6>.
- Hsu, Patrick D, David A Scott, Joshua A Weinstein, F Ann Ran, Silvana Konermann, Vineeta Agarwala, Yinqing Li, et al. 2013. "DNA Targeting Specificity of RNA-Guided Cas9 Nucleases." *Nature Biotechnology* 31 (9): 827–32. <https://doi.org/10.1038/nbt.2647>.
- Hu, Ruibo, Guang Qi, Yingzhen Kong, Dejing Kong, Qian Gao, and Gongke Zhou. 2010. "Comprehensive Analysis of NAC Domain Transcription Factor Gene Family in Populus Trichocarpa." *BMC Plant Biology* 10 (1): 145. <https://doi.org/10.1186/1471-2229-10-145>.
- Hu, Ruibo, Yan Xu, Changjiang Yu, Kang He, Qi Tang, Chunlin Jia, Guo He, Xiaoyu Wang, Yingzhen Kong, and Gongke Zhou. 2017. "Transcriptome Analysis of Genes Involved in Secondary Cell Wall Biosynthesis in Developing Internodes of Miscanthus Lutarioriparius." *Scientific Reports* 7 (1): 1–16. <https://doi.org/10.1038/s41598-017-08690-8>.
- Huang, Ji, Stefania Vendramin, Lizhen Shi, and Karen M. McGinnis. 2017. "Construction and Optimization of a Large Gene Coexpression Network in Maize Using RNA-Seq Data." *Plant Physiology* 175 (1): 568–83. <https://doi.org/10.1104/pp.17.00825>.
- Hughes, Thomas E., Jane A. Langdale, and Steven Kelly. 2014. "The Impact of Widespread Regulatory Neofunctionalization on Homeolog Gene Evolution Following Whole-Genome Duplication in Maize." *Genome Research* 24 (8): 1348–55.

<https://doi.org/10.1101/gr.172684.114>.

- Hussey, Steven G., Mohammed N. Saïdi, Charles A. Hefer, Alexander A. Myburg, and Jacqueline Grima-Pettenati. 2015. "Structural, Evolutionary and Functional Analysis of the NAC Domain Protein Family in Eucalyptus." *New Phytologist* 206 (4): 1337–50. <https://doi.org/10.1111/nph.13139>.
- Hussey, Steven G, Eshchar Mizrahi, Antanas V Spokevicius, Gerd Bossinger, Dave K Berger, and Alexander A Myburg. 2011. "SND2, a NAC Transcription Factor Gene, Regulates Genes Involved in Secondary Cell Wall Development in Arabidopsis Fibres and Increases Fibre Cell Area in Eucalyptus." *BMC Plant Biology* 11 (1): 173. <https://doi.org/10.1186/1471-2229-11-173>.
- Iiyama K, Lam TBT, Stone BA (1990) "Phenolic acid bridges between polysaccharides and lignin in wheat internodes". *Phytochemistry* 29:733-737
- Ishida, Yuji, Hideaki Saito, Shozo Ohta, Yukoh Hiei, Toshihiko Komari, and Takashi Kumashiro. 1996. "High Efficiency Transformation of Maize (*Zea Mays* L.) Mediated by *Agrobacterium Tumefaciens*." *Nature Biotechnology* 14 (6): 745–50. <https://doi.org/10.1038/nbt0696-745>.
- Ishii, Tadashi. 1997. "Structure and Functions of Feruloylated Polysaccharides." *Plant Science* 127 (2): 111–27. [https://doi.org/10.1016/S0168-9452\(97\)00130-1](https://doi.org/10.1016/S0168-9452(97)00130-1).
- Ismagul, Ainur, Nannan Yang, Elina Maltseva, Gulnur Iskakova, Inna Mazonka, Yuri Skiba, Huihui Bi, et al. 2018. "A Biolistic Method for High-Throughput Production of Transgenic Wheat Plants with Single Gene Insertions." *BMC Plant Biology* 18 (1): 1–8. <https://doi.org/10.1186/s12870-018-1326-1>.
- Ito, Masaki. 2005. "Conservation and Diversification of Three-Repeat Myb Transcription Factors in Plants." *Journal of Plant Research* 118 (1): 61–69. <https://doi.org/10.1007/s10265-005-0192-8>.
- Jain, Mukesh, Aashima Nijhawan, Rita Arora, Pinky Agarwal, Swatishmita Ray, Pooja Sharma, Sanjay Kapoor, Akhilesh K. Tyagi, and Jitendra P. Khurana. 2007. "F-Box Proteins in Rice. Genome-Wide Analysis, Classification, Temporal and Spatial Gene Expression during Panicle and Seed Development, and Regulation by Light and Abiotic Stress." *Plant Physiology* 143 (4): 1467–83. <https://doi.org/10.1104/pp.106.091900>.
- Jambagi, Shridhar, Rs Bhat, Sumangala Bhat, and Ms Kuruvinashetti. 2010. "Agrobacterium - Mediated Transformation Studies in Sorghum Using an Improved Gfp Reporter Gene." *SAT EJournal* 8 (December): 1–5.
- Janas, Krystyna M, Milena Cvikrová, Anna Pałagiewicz, and Josef Eder. 2000. "Alterations in Phenylpropanoid Content in Soybean Roots during Low Temperature Acclimation." *Plant Physiology and Biochemistry* 38 (7–8): 587–93. [https://doi.org/10.1016/S0981-9428\(00\)00778-6](https://doi.org/10.1016/S0981-9428(00)00778-6).
- Janda, Tibor, Gabriella Szalai, Krisztina Rios-Gonzalez, Ottó Veisz, and Emil Páldi. 2003. "Comparative Study of Frost Tolerance and Antioxidant Activity in Cereals." *Plant Science* 164 (2): 301–6. [https://doi.org/10.1016/S0168-9452\(02\)00414-4](https://doi.org/10.1016/S0168-9452(02)00414-4).
- Jarvis, Mike. 2003. "Cellulose Stacks Up." *Nature* 426 (6967): 611–12.

<https://doi.org/10.1038/426611a>.

- Jasin, M., and R. Rothstein. 2013. "Repair of Strand Breaks by Homologous Recombination." *Cold Spring Harbor Perspectives in Biology* 5 (11): a012740–a012740. <https://doi.org/10.1101/cshperspect.a012740>.
- Jawaid, Mohammad, M. T. Paridah, and Naheed Saba. 2017. *Introduction to Biomass and Its Composites. Lignocellulosic Fibre and Biomass-Based Composite Materials: Processing, Properties and Applications*. Elsevier Ltd. <https://doi.org/10.1016/B978-0-08-100959-8.00001-9>.
- Jensen, Jacob Krüger, Nathan Robert Johnson, and Curtis Gene Wilkerson. 2014. "Arabidopsis Thaliana IRX10 and Two Related Proteins from Psyllium and Physcomitrella Patens Are Xylan Xylosyltransferases." *The Plant Journal* 80 (2): 207–15. <https://doi.org/10.1111/tpj.12641>.
- Jeoung, J. M., S. Krishnaveni, S. Muthukrishnan, H. N. Trick, and G. H. Liang. 2002. "Optimization of Sorghum Transformation Parameters Using Genes for Green Fluorescent Protein and Beta-Glucuronidase as Visual Markers." *Hereditas* 137 (1): 20–28. <https://doi.org/10.1034/j.1601-5223.2002.1370104.x>.
- Jiang, Guo-Liang. 2015. "Molecular Marker-Assisted Breeding: A Plant Breeder's Review." In *Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools*, 431–72. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-22521-0_15.
- Jiao, Yinping, John J. Burke, Ratan Chopra, Gloria Burow, Junping Chen, Bo Wang, Chad Hayes, Yves Emendack, Doreen Ware, and Zhanguo Xin. 2016. "A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses." *The Plant Cell*, June, tpc.00373.2016. <https://doi.org/10.1105/tpc.16.00373>.
- Jin, H. 2000. "Transcriptional Repression by AtMYB4 Controls Production of UV-Protecting Sunscreens in Arabidopsis." *The EMBO Journal* 19 (22): 6150–61. <https://doi.org/10.1093/emboj/19.22.6150>.
- Jin, Jinpu, Feng Tian, De Chang Yang, Yu Qi Meng, Lei Kong, Jingchu Luo, and Ge Gao. 2017. "PlantTFDB 4.0: Toward a Central Hub for Transcription Factors and Regulatory Interactions in Plants." *Nucleic Acids Research* 45 (D1): D1040–45. <https://doi.org/10.1093/nar/gkw982>.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier. 2012. "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity." *Science* 337 (6096): 816–21. <https://doi.org/10.1126/science.1225829>.
- Joersbo, Morten, Lain Donaldson, Jette Kreiberg, Steen Guldager Petersen, Janne Brunstedt, and Finn T. Okkels. 1998. "Analysis of Mannose Selection Used for Transformation of Sugar Beet." *Molecular Breeding* 4 (2): 111–17. <https://doi.org/10.1023/A:1009633809610>.
- Jogeswar, G., D. Ranadheer, V. Anjaiah, and P. B. Kavi Kishor. 2007. "High Frequency Somatic Embryogenesis and Regeneration in Different Genotypes of Sorghum Bicolor (L.) Moench from Immature Inflorescence Explants." *In Vitro Cellular & Developmental Biology - Plant* 43 (2): 159–66. <https://doi.org/10.1007/s11627-007-9033-x>.

- Jørgensen, Kirsten, Anne Vinther Rasmussen, Marc Morant, Allan Holm Nielsen, Nanna Bjarnholt, Mika Zagrobelny, Søren Bak, and Birger Lindberg Møller. 2005. "Metabolon Formation and Metabolic Channeling in the Biosynthesis of Plant Natural Products." *Current Opinion in Plant Biology* 8 (3): 280–91. <https://doi.org/10.1016/j.pbi.2005.03.014>.
- Joshi, Chandrashekhar P., Shivegowda Thammannagowda, Takeshi Fujino, Ji-Qing Gou, Utku Avci, Candace H. Haigler, Lisa M. McDonnell, et al. 2011. "Perturbation of Wood Cellulose Synthesis Causes Pleiotropic Effects in Transgenic Aspen." *Molecular Plant* 4 (2): 331–45. <https://doi.org/10.1093/mp/ssq081>.
- Kadier, Yibadaiti, Yi yi Zu, Qing min Dai, Ge Song, Shi wen Lin, Qing peng Sun, Jin bao Pan, and Min Lu. 2017. "Genome-Wide Identification, Classification and Expression Analysis of NAC Family of Genes in Sorghum [Sorghum Bicolor (L.) Moench]." *Plant Growth Regulation* 83 (2): 301–12. <https://doi.org/10.1007/s10725-017-0295-y>.
- Kaeppler, H. F., and J. F. Pedersen. 1996. "Media Effects on Phenotype of Callus Cultures Initiated from Photoperiod-Insensitive, Elite Inbred Sorghum Lines." *Maydica* 41 (2): 83–89.
- Kapila, Jyoti, Riet De Rycke, Marc Van Montagu, and Geert Angenon. 1997. "An Agrobacterium-Mediated Transient Gene Expression System for Intact Leaves." *Plant Science* 122 (1): 101–8. [https://doi.org/10.1016/S0168-9452\(96\)04541-4](https://doi.org/10.1016/S0168-9452(96)04541-4).
- Katiyar, Amit, Shuchi Smita, Sangram K. Lenka, Ravi Rajwanshi, Viswanathan Chinnusamy, and Kailash C. Bansal. 2012. "Genome-Wide Classification and Expression Analysis of MYB Transcription Factor Families in Rice and Arabidopsis." *BMC Genomics* 13 (1): 1. <https://doi.org/10.1186/1471-2164-13-544>.
- Katoh, Kazutaka, John Rozewicki, and Kazunori D. Yamada. 2017. "MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization." *Briefings in Bioinformatics*, no. June: 1–7. <https://doi.org/10.1093/bib/bbx108>.
- Kayal, W. El. 2006. "Expression Profile of CBF-like Transcriptional Factor Genes from Eucalyptus in Response to Cold." *Journal of Experimental Botany* 57 (10): 2455–69. <https://doi.org/10.1093/jxb/erl019>.
- Kebrom, Tesfamichael H., Brian McKinley, and John E. Mullet. 2017. "Dynamics of Gene Expression during Development and Expansion of Vegetative Stem Internodes of Bioenergy Sorghum." *Biotechnology for Biofuels* 10 (1): 1–16. <https://doi.org/10.1186/s13068-017-0848-3>.
- Kende, Hans, Esther van der Knaap, and Hyung-Taeg Cho. 1998. "Deepwater Rice: A Model Plant to Study Stem Elongation." *Plant Physiology* 118 (4): 1105–10. <https://doi.org/10.1104/pp.118.4.1105>.
- Kerstetter, Randall A., Krista Bollman, R. Alexandra Taylor, Kirsten Bomblies, and R. Scott Poethig. 2001. "KANADI Regulates Organ Polarity in Arabidopsis." *Nature* 411 (6838): 706–9. <https://doi.org/10.1038/35079629>.
- Keshavareddy, G., A.R.V. Kumar, and Vemanna S. Ramu. 2018. "Methods of Plant Transformation-

- A Review.” *International Journal of Current Microbiology and Applied Sciences* 7 (07): 2656–68. <https://doi.org/10.20546/ijcmas.2018.707.312>.
- Kieffer, F. 2000. “The Fungal Elicitor Cryptogein Induces Cell Wall Modifications on Tobacco Cell Suspension.” *Journal of Experimental Botany* 51 (352): 1799–1811. <https://doi.org/10.1093/jexbot/51.352.1799>.
- Kikkert, J.R., G.S. Ali, P.G. Wallace, B. Reisch, and G.M. Reustle. 2000. “Expression Of A Fungal Chitinase In *Vitis Vinifera* L. ‘Merlot’ And ‘Chardonnay’ Plants Produced By Biolistic Transformation.” *Acta Horticulturae*, no. 528 (May): 299–306. <https://doi.org/10.17660/ActaHortic.2000.528.41>.
- Kim, Daehwan, Ben Langmead, and Steven L Salzberg. 2015. “HISAT: A Fast Spliced Aligner with Low Memory Requirements.” *Nature Methods* 12 (March): 357. <https://doi.org/10.1038/nmeth.3317>.
- Kim, Hae Koo, Delphine Luquet, Erik van Oosterom, Michael Dingkuhn, and Graeme Hammer. 2010. “Regulation of Tillering in Sorghum: Genotypic Effects.” *Annals of Botany* 106 (1): 69–78. <https://doi.org/10.1093/aob/mcq080>.
- Kim, Misook, and Donal F. Day. 2011. “Composition of Sugar Cane, Energy Cane, and Sweet Sorghum Suitable for Ethanol Production at Louisiana Sugar Mills.” *Journal of Industrial Microbiology and Biotechnology* 38 (7): 803–7. <https://doi.org/10.1007/s10295-010-0812-8>.
- Kim, Won Chan, Joo Yeol Kim, Jae Heung Ko, Jungmook Kim, and Kyung Hwan Han. 2013. “Transcription Factor MYB46 Is an Obligate Component of the Transcriptional Regulatory Complex for Functional Expression of Secondary Wall-Associated Cellulose Synthases in *Arabidopsis Thaliana*.” *Journal of Plant Physiology* 170 (15): 1374–78. <https://doi.org/10.1016/j.jplph.2013.04.012>.
- Kim, Youn-Sung, Sang-Gyu Kim, Jung-Eun Park, Hye-Young Park, Mi-Hye Lim, Nam-Hai Chua, and Chung-Mo Park. 2006. “A Membrane-Bound NAC Transcription Factor Regulates Cell Division in *Arabidopsis*.” *The Plant Cell* 18 (11): 3132–44. <https://doi.org/10.1105/tpc.106.043018>.
- Kitajima, Aya, Satoru Asatsuma, Hisao Okada, Yuki Hamada, Kentaro Kaneko, Yohei Nanjo, Yasushi Kawagoe, et al. 2009. “The Rice α -Amylase Glycoprotein Is Targeted from the Golgi Apparatus through the Secretory Pathway to the Plastids.” *The Plant Cell* 21 (9): 2844–58. <https://doi.org/10.1105/tpc.109.068288>.
- Ko, J. H., H. W. Jeon, W. C. Kim, J. Y. Kim, and K. H. Han. 2014. “The MYB46/MYB83-Mediated Transcriptional Regulatory Programme Is a Gatekeeper of Secondary Wall Biosynthesis.” *Annals of Botany* 114 (6): 1099–1107. <https://doi.org/10.1093/aob/mcu126>.
- Ko, Jae-Heung, Won-Chan Kim, and Kyung-Hwan Han. 2009. “Ectopic Expression of MYB46 Identifies Transcriptional Regulatory Genes Involved in Secondary Wall Biosynthesis in *Arabidopsis*.” *The Plant Journal* 60 (4): 649–65. <https://doi.org/10.1111/j.1365->

- Koetle, M.J., J.F. Finnie, E. Balázs, and J. Van Staden. 2015. "A Review on Factors Affecting the Agrobacterium -Mediated Genetic Transformation in Ornamental Monocotyledonous Geophytes." *South African Journal of Botany* 98 (May): 37–44. <https://doi.org/10.1016/j.sajb.2015.02.001>.
- König, Stefanie, Kirstin Feussner, Alexander Kaever, Manuel Landesfeind, Corinna Thurow, Petr Karlovsky, Christiane Gatz, Andrea Polle, and Ivo Feussner. 2014. "Soluble Phenylpropanoids Are Involved in the Defense Response of Arabidopsis against Verticillium Longisporum." *New Phytologist* 202 (3): 823–37. <https://doi.org/10.1111/nph.12709>.
- Krens, F. A., L. Molendijk, G. J. Wullems, and R. A. Schilperoort. 1982. "In Vitro Transformation of Plant Protoplasts with Ti-Plasmid DNA." *Nature* 296 (5852): 72–74. <https://doi.org/10.1038/296072a0>.
- Krouk, Gabriel, Jesse Lingeman, Amy Marshall Colon, Gloria Coruzzi, and Dennis Shasha. 2013. "Gene Regulatory Networks in Plants: Learning Causality from Time and Perturbation." *Genome Biology* 14 (6): 123. <https://doi.org/10.1186/gb-2013-14-6-123>.
- Kruijer, Willem, Martin P. Boer, Marcos Malosetti, Pádraic J. Flood, Bas Engel, Rik Kooke, Joost J. B. Keurentjes, and Fred A. van Eeuwijk. 2015. "Marker-Based Estimation of Heritability in Immortal Populations." *Genetics* 199 (2): 379–98. <https://doi.org/10.1534/genetics.114.167916>.
- Kubo, M. 2005. "Transcription Switches for Protoxylem and Metaxylem Vessel Formation." *Genes & Development* 19 (16): 1855–60. <https://doi.org/10.1101/gad.1331305>.
- Kumar, Vinod, LeAnne M. Campbell, and Keerti S. Rathore. 2011. "Rapid Recovery- and Characterization of Transformants Following Agrobacterium-Mediated T-DNA Transfer to Sorghum." *Plant Cell, Tissue and Organ Culture (PCTOC)* 104 (2): 137–46. <https://doi.org/10.1007/s11240-010-9809-2>.
- Kuriyama, Tomoko, Setsuko Shimada, and Minami Matsui. 2019. "Improvement of Agrobacterium-Mediated Transformation for Tannin-Producing Sorghum." *Plant Biotechnology* 36 (1): 43–48. <https://doi.org/10.5511/plantbiotechnology.19.0131a>.
- Kyozuka, Junko, Yasuyuki Hayashi, and Ko Shimamoto. 1987. "High Frequency Plant Regeneration from Rice Protoplasts by Novel Nurse Culture Methods." *MGG Molecular & General Genetics* 206 (3): 408–13. <https://doi.org/10.1007/BF00428879>.
- Lam, Pui Ying, Yuki Tobimatsu, Yuri Takeda, Shiro Suzuki, Masaomi Yamamura, Toshiaki Umezawa, and Clive Lo. 2017. "Disrupting Flavone Synthase II Alters Lignin and Improves Biomass Digestibility." *Plant Physiology* 174 (2): 972–85. <https://doi.org/10.1104/pp.16.01973>.
- Lam, Sheung Kwan, Ching Lung Siu, Stefan Hillmer, Seonghoe Jang, Gynheung An, David G. Robinson, and Liwen Jiang. 2007. "Rice SCAMP1 Defines Clathrin-Coated, Trans -Golgi-Located Tubular-Vesicular Structures as an Early Endosome in Tobacco BY-2 Cells." *The Plant Cell* 19 (1): 296–319. <https://doi.org/10.1105/tpc.106.045708>.

- Langfelder, Peter, and Steve Horvath. 2008. "WGCNA: An R Package for Weighted Correlation Network Analysis." *BMC Bioinformatics* 9. <https://doi.org/10.1186/1471-2105-9-559>.
- Langfelder, Peter, and Steve Horvath. 2012. "Fast R Functions for Robust Correlations and Hierarchical Clustering." *Journal of Statistical Software* 46 (11). <https://doi.org/10.18637/jss.v046.i11>.
- Lee, Sangmin, Pil Joon Seo, Hyo-Jun Lee, and Chung-Mo Park. 2012. "A NAC Transcription Factor NTL4 Promotes Reactive Oxygen Species Production during Drought-Induced Leaf Senescence in Arabidopsis." *The Plant Journal* 70 (5): 831–44. <https://doi.org/10.1111/j.1365-313X.2012.04932.x>.
- Leene, Jelle Van, Jonas Blomme, Shubhada R Kulkarni, Bernard Cannoot, Nancy De Winne, Dominique Eeckhout, Geert Persiau, et al. 2016. "Functional Characterization of the Arabidopsis Transcription Factor BZIP29 Reveals Its Role in Leaf and Root Development." *Journal of Experimental Botany* 67 (19): 5825–40. <https://doi.org/10.1093/jxb/erw347>.
- Legay, Sylvain, Pierre Sivadon, Anne Sophie Blervacq, Nathalie Pavy, Ahmad Baghdady, Laurence Tremblay, Caroline Levasseur, et al. 2010. "EgMYB1, an R2R3 MYB Transcription Factor from Eucalyptus Negatively Regulates Secondary Cell Wall Formation in Arabidopsis and Poplar." *New Phytologist* 188 (3): 774–86. <https://doi.org/10.1111/j.1469-8137.2010.03432.x>.
- Leucci, Maria Rosaria, Marcello Salvatore Lenucci, Gabriella Piro, and Giuseppe Dalessandro. 2008. "Water Stress and Cell Wall Polysaccharides in the Apical Root Zone of Wheat Cultivars Varying in Drought Tolerance." *Journal of Plant Physiology* 165 (11): 1168–80. <https://doi.org/10.1016/j.jplph.2007.09.006>.
- Li, Aixia, Shangang Jia, Abou Yobi, Zhengxiang Ge, Shirley J. Sato, Chi Zhang, Ruthie Angelovici, Thomas E. Clemente, and David R. Holding. 2018. "Editing of an Alpha-Kafirin Gene Family Increases, Digestibility and Protein Quality in Sorghum." *Plant Physiology* 177 (4): 1425–38. <https://doi.org/10.1104/pp.18.00200>.
- Li, Eryang, Apurva Bhargava, Weiya Qiang, Michael C. Friedmann, Natascha Forneris, Rodney A. Savidge, Lee A. Johnson, Shawn D. Mansfield, Brian E. Ellis, and Carl J. Douglas. 2012. "The Class II KNOX Gene KNAT7 Negatively Regulates Secondary Wall Formation in Arabidopsis and Is Functionally Conserved in Populus." *New Phytologist* 194 (1): 102–15. <https://doi.org/10.1111/j.1469-8137.2011.04016.x>.
- Li, Q., Y.-C. Lin, Y.-H. Sun, J. Song, H. Chen, X.-H. Zhang, R. R. Sederoff, and V. L. Chiang. 2012. "Splice Variant of the SND1 Transcription Factor Is a Dominant Negative of SND1 Members and Their Regulation in Populus Trichocarpa." *Proceedings of the National Academy of Sciences* 109 (36): 14699–704. <https://doi.org/10.1073/pnas.1212977109>.
- Li, Xiaoxu, Cun Guo, Salman Ahmad, Qi Wang, Jing Yu, and Cheng Liu. 2019. "Systematic Analysis of MYB Family Genes in Potato and Their Multiple Roles in Development and Stress Responses," 1–21.

- Li, Zhenjun, Rihe Peng, Yongsheng Tian, Hongjuan Han, Jing Xu, and Quanhong Yao. 2016. “Genome-Wide Identification and Analysis of the MYB Transcription Factor Superfamily in *Solanum Lycopersicum*.” *Plant and Cell Physiology* 57 (8): 1657–77. <https://doi.org/10.1093/pcp/pcw091>.
- Liepman, Aaron H., Curtis G. Wilkerson, and Kenneth Keegstra. 2005. “Expression of Cellulose Synthase-like (Csl) Genes in Insect Cells Reveals That CslA Family Members Encode Mannan Synthases.” *Proceedings of the National Academy of Sciences of the United States of America* 102 (6): 2221–26. <https://doi.org/10.1073/pnas.0409179102>.
- Lin, Hung ying, Qiang Liu, Xiao Li, Jinliang Yang, Sanzhen Liu, Yinlian Huang, Michael J. Scanlon, Dan Nettleton, and Patrick S. Schnable. 2017. “Substantial Contribution of Genetic Variation in the Expression of Transcription Factors to Phenotypic Variation Revealed by ERD-GWAS.” *Genome Biology* 18 (1): 1–14. <https://doi.org/10.1186/s13059-017-1328-6>.
- Lin, Yan, and Shuzo Tanaka. 2006. “Ethanol Fermentation from Biomass Resources: Current State and Prospects.” *Applied Microbiology and Biotechnology* 69 (6): 627–42. <https://doi.org/10.1007/s00253-005-0229-x>.
- Lin, Ying-Chung Jimmy, Hao Chen, Quanzi Li, Wei Li, Jack P. Wang, Rui Shi, Sermsawat Tunlaya-Anukit, et al. 2017. “Reciprocal Cross-Regulation of VND and SND Multigene TF Families for Wood Formation in *Populus Trichocarpa*.” *Proceedings of the National Academy of Sciences* 114 (45): E9722–29. <https://doi.org/10.1073/pnas.1714422114>.
- Lindsay, Shona E., and Stephen C. Fry. 2007. “Control of Diferulate Formation in Dicotyledonous and Gramineous Cell-Suspension Cultures.” *Planta* 227 (2): 439–52. <https://doi.org/10.1007/s00425-007-0630-z>.
- Ling, Lei, Lili Song, Youjing Wang, and Changhong Guo. 2017. “Genome-Wide Analysis and Expression Patterns of the NAC Transcription Factor Family in *Medicago Truncatula*.” *Physiology and Molecular Biology of Plants* 23 (2): 343–56. <https://doi.org/10.1007/s12298-017-0421-3>.
- Lionetti, V., F. Francocci, S. Ferrari, C. Volpi, D. Bellincampi, R. Galletti, R. D’Ovidio, G. De Lorenzo, and F. Cervone. 2010. “Engineering the Cell Wall by Reducing De-Methyl-Esterified Homogalacturonan Improves Saccharification of Plant Tissues for Bioconversion.” *Proceedings of the National Academy of Sciences* 107 (2): 616–21. <https://doi.org/10.1073/pnas.0907549107>.
- Liu, Guoquan, Bradley C. Campbell, and Ian D. Godwin. 2014. “Sorghum Genetic Transformation by Particle Bombardment.” In , 219–34. https://doi.org/10.1007/978-1-62703-715-0_18.
- Liu, Guoquan, and Ian D. Godwin. 2012. “Highly Efficient Sorghum Transformation.” *Plant Cell Reports* 31 (6): 999–1007. <https://doi.org/10.1007/s00299-011-1218-4>.
- Liu, Guoquan, Jieqing Li, and Ian D. Godwin. 2019. “Genome Editing by CRISPR/Cas9 in Sorghum Through Biolistic Bombardment.” In , 169–83. https://doi.org/10.1007/978-1-4939-9039-9_12.
- Liu, Wei, Liping Lin, Zhiyuan Zhang, Siqi Liu, Kuan Gao, Yanbin Lv, Huan Tao, and Huaqin He.

2019. “Gene Co-Expression Network Analysis Identifies Trait-Related Modules in *Arabidopsis Thaliana*.” *Planta* 249 (5): 1487–1501. <https://doi.org/10.1007/s00425-019-03102-9>.
- Liu, Yuanyuan, and Carl J Douglas. 2015. “A Role for OVATE FAMILY PROTEIN1 (OFP1) and OFP4 in a BLH6-KNAT7 Multi-Protein Complex Regulating Secondary Cell Wall Formation in *Arabidopsis Thaliana*.” *Plant Signaling & Behavior* 10 (7): e1033126. <https://doi.org/10.1080/15592324.2015.1033126>.
- Liu, Yuanyuan, Shijun You, Mallorie Taylor-Teeple, Wenhua L. Li, Mathias Schuetz, Siobhan M. Brady, and Carl J. Douglas. 2014. “BEL1-LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 Interact and Regulate Secondary Cell Wall Formation via Repression of *REVOLUTA*.” *The Plant Cell Online* 26 (12): 4843–61. <https://doi.org/10.1105/tpc.114.128322>.
- Loix, Christophe, Michiel Huybrechts, Jaco Vangronsveld, Marijke Gielen, Els Keunen, and Ann Cuypers. 2017. “Reciprocal Interactions between Cadmium-Induced Cell Wall Responses and Oxidative Stress in Plants.” *Frontiers in Plant Science* 8 (October). <https://doi.org/10.3389/fpls.2017.01867>.
- Lourenço, Ana, and Helena Pereira. 2018. “Compositional Variability of Lignin in Biomass.” In *Lignin - Trends and Applications*. InTech. <https://doi.org/10.5772/intechopen.71208>.
- Lu, Sheen X., Stephen M. Knowles, Christos Andronis, May S. Ong, and Elaine M. Tobin. 2009. “CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL Function Synergistically in the Circadian Clock of *Arabidopsis*.” *Plant Physiology* 150 (2): 834–43. <https://doi.org/10.1104/pp.108.133272>.
- M. Martin. 2011. “Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads.” *EMBnet J.* 17 (1): 10–12. <http://journal.embnnet.org/index.php/embnnetjournal/article/view/200/458%5Cnpapers3://publication/uuid/5AF8D3EB-3B82-42AB-82ED-8B1B32092FFC>.
- Ma, Xingliang, Qunyu Zhang, Qinlong Zhu, Wei Liu, Yan Chen, Rong Qiu, Bin Wang, et al. 2015. “A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants.” *Molecular Plant* 8 (8): 1274–84. <https://doi.org/10.1016/j.molp.2015.04.007>.
- Ma, Xuelian, Hansheng Zhao, Wenying Xu, Qi You, Hengyu Yan, Zhimin Gao, and Zhen Su. 2018. “Co-Expression Gene Network Analysis and Functional Module Identification in Bamboo Growth and Development.” *Frontiers in Genetics* 9 (November): 1–15. <https://doi.org/10.3389/fgene.2018.00574>.
- Maas, C., C. Reichel, J. Schell, and H. H. Steinbiß. 1995. “Preparation and Transformation of Monocot Protoplasts.” *Methods in Cell Biology* 50 (C): 383–99. [https://doi.org/10.1016/S0091-679X\(08\)61045-6](https://doi.org/10.1016/S0091-679X(08)61045-6).
- Maćkowska, Katarzyna, Agata Jarosz, and Ewa Grzebelus. 2014. “Plant Regeneration from Leaf-

- Derived Protoplasts within the *Daucus* Genus: Effect of Different Conditions in Alginate Embedding and Phytosulfokine Application.” *Plant Cell, Tissue and Organ Culture (PCTOC)* 117 (2): 241–52. <https://doi.org/10.1007/s11240-014-0436-1>.
- Maia, Josemir M., Eduardo L. Voigt, Sérgio L. Ferreira-Silva, Adilton de V. Fontenele, Cristiane E. C. Macêdo, and Joaquim A. G. Silveira. 2013. “Differences in Cowpea Root Growth Triggered by Salinity and Dehydration Are Associated with Oxidative Modulation Involving Types I and III Peroxidases and Apoplastic Ascorbate.” *Journal of Plant Growth Regulation* 32 (2): 376–87. <https://doi.org/10.1007/s00344-012-9308-2>.
- Maleki, Samaneh Sadat, Kourosh Mohammadi, and Kong Shu Ji. 2016. “Characterization of Cellulose Synthesis in Plant Cells.” *Scientific World Journal* 2016. <https://doi.org/10.1155/2016/8641373>.
- Mann, David GJ, Zachary R King, Wusheng Liu, Blake L Joyce, Ryan J Percifield, Jennifer S Hawkins, Peter R LaFayette, et al. 2011. “Switchgrass (*Panicum Virgatum* L.) Polyubiquitin Gene (PvUbi1 and PvUbi2) Promoters for Use in Plant Transformation.” *BMC Biotechnology* 11 (1): 74. <https://doi.org/10.1186/1472-6750-11-74>.
- Manova, Vasilissa, and Damian Gruszka. 2015. “DNA Damage and Repair in Plants – from Models to Crops.” *Frontiers in Plant Science* 6 (October). <https://doi.org/10.3389/fpls.2015.00885>.
- Marinova, Krasimira, Lucille Pourcel, Barbara Weder, Michael Schwarz, Denis Barron, Jean-Marc Routaboul, Isabelle Debeaujon, and Markus Klein. 2007. “The Arabidopsis MATE Transporter TT12 Acts as a Vacuolar Flavonoid/H⁺ -Antiporter Active in Proanthocyanidin-Accumulating Cells of the Seed Coat.” *The Plant Cell* 19 (6): 2023–38. <https://doi.org/10.1105/tpc.106.046029>.
- Masani, Mat Yunus Abdul, Gundula A. Noll, Ghulam Kadir Ahmad Parveez, Ravigadevi Sambanthamurthi, and Dirk Prüfer. 2014. “Efficient Transformation of Oil Palm Protoplasts by PEG-Mediated Transfection and DNA Microinjection.” Edited by Meng-xiang Sun. *PLoS ONE* 9 (5): e96831. <https://doi.org/10.1371/journal.pone.0096831>.
- Matsui, Kyoko, Yoshimi Umemura, and Masaru Ohme-Takagi. 2008. “AtMYBL2, a Protein with a Single MYB Domain, Acts as a Negative Regulator of Anthocyanin Biosynthesis in Arabidopsis.” *The Plant Journal* 55 (6): 954–67. <https://doi.org/10.1111/j.1365-313X.2008.03565.x>.
- McBee, George G., and F. R. Miller. 1982. “Carbohydrates in Sorghum Culms as Influenced by Cultivars, Spacing, and Maturity over a Diurnal Period 1.” *Crop Science* 22 (2): 381–85. <https://doi.org/10.2135/cropsci1982.0011183X002200020042x>.
- McCarthy, Ryan L., Ruiqin Zhong, Scott Fowler, David Lyskowski, Hasanthi Piyasena, Kara Carleton, Charles Spicer, and Zheng-Hua Ye. 2010. “The Poplar MYB Transcription Factors, PtrMYB3 and PtrMYB20, Are Involved in the Regulation of Secondary Wall Biosynthesis.” *Plant and Cell Physiology* 51 (6): 1084–90. <https://doi.org/10.1093/pcp/pcq064>.
- McCarthy, R. L., R. Zhong, and Z.-H. Ye. 2009. “MYB83 Is a Direct Target of SND1 and Acts Redundantly with MYB46 in the Regulation of Secondary Cell Wall Biosynthesis in

- Arabidopsis.” *Plant and Cell Physiology* 50 (11): 1950–64. <https://doi.org/10.1093/pcp/pcp139>.
- McCormick, Ryan F., Sandra K. Truong, Avinash Sreedasyam, Jerry Jenkins, Shengqiang Shu, David Sims, Megan Kennedy, et al. 2018. “The Sorghum Bicolor Reference Genome: Improved Assembly, Gene Annotations, a Transcriptome Atlas, and Signatures of Genome Organization.” *Plant Journal* 93 (2): 338–54. <https://doi.org/10.1111/tpj.13781>.
- McFarlane, Heather E., Anett Döring, and Staffan Persson. 2014. “The Cell Biology of Cellulose Synthesis.” *Annual Review of Plant Biology* 65 (1): 69–94. <https://doi.org/10.1146/annurev-arplant-050213-040240>.
- McKinley, Brian, William Rooney, Curtis Wilkerson, and John Mullet. 2016. “Dynamics of Biomass Partitioning, Stem Gene Expression, Cell Wall Biosynthesis, and Sucrose Accumulation during Development of *Sorghum Bicolor*.” *The Plant Journal*, 1–19. <https://doi.org/10.1111/tpj.13269>.
- Meents, Miranda J, Yoichiro Watanabe, and A Lacey Samuels. 2018. “The Cell Biology of Secondary Cell Wall Biosynthesis.” *Annals of Botany* 121 (6): 1107–25. <https://doi.org/10.1093/aob/mcy005>.
- Mellerowicz, Ewa J., Marie Baucher, Björn Sundberg, and Wout Boerjan. 2001. “Unravelling Cell Wall Formation in the Woody Dicot Stem.” *Plant Molecular Biology* 47 (1–2): 239–74. <https://doi.org/10.1023/A:1010699919325>.
- Miao, Jin, Dongshu Guo, Jinzhe Zhang, Qingpei Huang, Genji Qin, Xin Zhang, Jianmin Wan, Hongya Gu, and Li-Jia Qu. 2013. “Targeted Mutagenesis in Rice Using CRISPR-Cas System.” *Cell Research* 23 (10): 1233–36. <https://doi.org/10.1038/cr.2013.123>.
- Miedes, Eva, Ruben Vanholme, Wout Boerjan, and Antonio Molina. 2014. “The Role of the Secondary Cell Wall in Plant Resistance to Pathogens.” *Frontiers in Plant Science* 5 (AUG): 1–13. <https://doi.org/10.3389/fpls.2014.00358>.
- Mierziak, Justyna, Kamil Kostyn, and Anna Kulma. 2014. “Flavonoids as Important Molecules of Plant Interactions with the Environment.” *Molecules* 19 (10): 16240–65. <https://doi.org/10.3390/molecules191016240>.
- Miki, Brian, and Sylvia McHugh. 2004. “Selectable Marker Genes in Transgenic Plants: Applications, Alternatives and Biosafety.” *Journal of Biotechnology* 107 (3): 193–232. <https://doi.org/10.1016/j.jbiotec.2003.10.011>.
- Miller, F.R. 1984. Registration of RTx430 sorghum parental line. Crop the experimental line. Since that time, these lines have been Sci. 24:1224
- Miller, Gad, Nobuhiro Suzuki, Sultan Ciftci-Yilmaz, and Ron Mittler. 2010. “Reactive Oxygen Species Homeostasis and Signalling during Drought and Salinity Stresses.” *Plant, Cell & Environment* 33 (4): 453–67. <https://doi.org/10.1111/j.1365-3040.2009.02041.x>.
- Mitchell, Rowan A.C., Paul Dupree, and Peter R. Shewry. 2007. “A Novel Bioinformatics Approach Identifies Candidate Genes for the Synthesis and Feruloylation of Arabinoxylan.” *Plant Physiology* 144 (1): 43–53. <https://doi.org/10.1104/pp.106.094995>.

- Mitsuda, Nobutaka, Akira Iwase, Hiroyuki Yamamoto, Masato Yoshida, Motoaki Seki, Kazuo Shinozaki, and Masaru Ohme-Takagi. 2007. "NAC Transcription Factors, NST1 and NST3, Are Key Regulators of the Formation of Secondary Walls in Woody Tissues of Arabidopsis." *The Plant Cell Online* 19 (1): 270–80. <https://doi.org/10.1105/tpc.106.047043>.
- Mitsuda, Nobutaka, Motoaki Seki, Kazuo Shinozaki, and Masaru Ohme-takagi. 2005. "The NAC Transcription Factors NST1 and NST2 of Arabidopsis Regulate Secondary Wall Thickenings and Are Required for Anther Dehiscence." *The Plant Cell* 17 (November): 2993–3006. <https://doi.org/10.1105/tpc.105.036004.1>.
- Mliki, A., R. Jardak, G. M. Reustle, and A. Ghorbel. 2003. "Isolation and Culture of Leaf Protoplasts from Tunisian Grapes." *Journal International Des Sciences de La Vigne et Du Vin* 37 (3): 145–53. <https://doi.org/10.20870/oeno-one.2003.37.3.952>.
- Mohnen, D. 2008. "Pectin Structure and Biosynthesis." *Current Opinion in Plant Biology* 11 (3): 266–77. <https://doi.org/10.1016/j.pbi.2008.03.006>.
- Mohr, Alison, and Sujatha Raman. 2013. "Lessons from First Generation Biofuels and Implications for the Sustainability Appraisal of Second Generation Biofuels." *Energy Policy* 63 (December): 114–22. <https://doi.org/10.1016/j.enpol.2013.08.033>.
- Mølhøj, Michael, Silvere Pagant, and Herman Höfte. 2002. "Towards Understanding the Role of Membrane-Bound Endo- β -1,4-Glucanases in Cellulose Biosynthesis." *Plant and Cell Physiology* 43 (12): 1399–1406. <https://doi.org/10.1093/pcp/pcf163>.
- Molinari, Hugo B. C., Till K. Pellny, Jackie Freeman, Peter R. Shewry, and Rowan A. C. Mitchell. 2013. "Grass Cell Wall Feruloylation: Distribution of Bound Ferulate and Candidate Gene Expression in Brachypodium Distachyon." *Frontiers in Plant Science* 4. <https://doi.org/10.3389/fpls.2013.00050>.
- Monk, R.L., F.R. Miller, and G.G. McBee. 1984. "Sorghum Improvement for Energy Production." *Biomass* 6 (1–2): 145–53. [https://doi.org/10.1016/0144-4565\(84\)90017-9](https://doi.org/10.1016/0144-4565(84)90017-9).
- Mookkan, Muruganantham, Kimberly Nelson-Vasilchik, Joel Hague, Zhanyuan J. Zhang, and Albert P. Kausch. 2017. "Selectable Marker Independent Transformation of Recalcitrant Maize Inbred B73 and Sorghum P898012 Mediated by Morphogenic Regulators BABY BOOM and WUSCHEL2." *Plant Cell Reports* 36 (9): 1477–91. <https://doi.org/10.1007/s00299-017-2169-1>.
- Moscou, M. J., and A. J. Bogdanove. 2009. "A Simple Cipher Governs DNA Recognition by TAL Effectors." *Science* 326 (5959): 1501–1501. <https://doi.org/10.1126/science.1178817>.
- Moura, Jullyana Cristina Magalhães Silva, Cesar Augusto Valencise Bonine, Juliana de Oliveira Fernandes Viana, Marcelo Carnier Dornelas, and Paulo Mazzafera. 2010. "Abiotic and Biotic Stresses and Changes in the Lignin Content and Composition in Plants." *Journal of Integrative Plant Biology* 52 (4): 360–76. <https://doi.org/10.1111/j.1744-7909.2010.00892.x>.
- Mu, Guiqin, Naidi Chang, Kui Xiang, Yaou Sheng, Zhiming Zhang, and Guangtang Pan. 2012. "Genetic Transformation of Maize Female Inflorescence Following Floral Dip Method Mediated

- by *Agrobacterium*.” *Biotechnology(Faisalabad)* 11 (3): 178–83.
<https://doi.org/10.3923/biotech.2012.178.183>.
- Murakami, Y. 1999. “The Differentiation of Contact Cells and Isolation Cells in the Xylem Ray Parenchyma of *Populus Maximowiczii*.” *Annals of Botany* 84 (4): 429–35.
<https://doi.org/10.1006/anbo.1999.0931>.
- Murashige, Toshio, and Folke Skoog. 1962. “A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures.” *Physiologia Plantarum* 15 (3): 473–97.
<https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>.
- Murray, Seth C., William L. Rooney, Sharon E. Mitchell, Arun Sharma, Patricia E. Klein, John E. Mullet, and Stephen Kresovich. 2008. “Genetic Improvement of Sorghum as a Biofuel Feedstock: II. QTL for Stem and Leaf Structural Carbohydrates.” *Crop Science* 48 (6): 2180–93.
<https://doi.org/10.2135/cropsci2008.01.0068>.
- Murray, Seth C., Arun Sharma, William L. Rooney, Patricia E. Klein, John E. Mullet, Sharon E. Mitchell, and Stephen Kresovich. 2008. “Genetic Improvement of Sorghum as a Biofuel Feedstock: I. QTL for Stem Sugar and Grain Nonstructural Carbohydrates.” *Crop Science* 48 (6): 2165–79. <https://doi.org/10.2135/cropsci2008.01.0016>.
- Muthamilarasan, Mehanathan, Rohit Khandelwal, Chandra Bhan Yadav, Venkata Suresh Bonthala, Yusuf Khan, and Manoj Prasad. 2014. “Identification and Molecular Characterization of MYB Transcription Factor Superfamily in C 4 Model Plant Foxtail Millet (*Setaria Italica* L .)” 9 (10).
<https://doi.org/10.1371/journal.pone.0109920>.
- Nakano, Yoshimi, Masatoshi Yamaguchi, Hitoshi Endo, Nur Ardiyana Rejab, and Misato Ohtani. 2015. “NAC-MYB-Based Transcriptional Regulation of Secondary Cell Wall Biosynthesis in Land Plants.” *Frontiers in Plant Science* 6 (May): 1–18. <https://doi.org/10.3389/fpls.2015.00288>.
- Nakashima, Kazuo, Kazuko Yamaguchi-Shinozaki, and Kazuo Shinozaki. 2014. “The Transcriptional Regulatory Network in the Drought Response and Its Crosstalk in Abiotic Stress Responses Including Drought, Cold, and Heat.” *Frontiers in Plant Science* 5 (May).
<https://doi.org/10.3389/fpls.2014.00170>.
- Naoumkina, Marina A., Qiao Zhao, Lina Gallego-Giraldo, Xinbin Dai, Patrick X. Zhao, and Richard A. Dixon. 2010. “Genome-Wide Analysis of Phenylpropanoid Defence Pathways.” *Molecular Plant Pathology*, July, no-no. <https://doi.org/10.1111/j.1364-3703.2010.00648.x>.
- Nasidi, Muhammad, Reginald Agu, Yusuf Deeni, and Graeme Walker. 2015. “Improved Production of Ethanol Using Bagasse from Different Sorghum Cultivars.” *Biomass and Bioenergy* 72 (January): 288–99. <https://doi.org/10.1016/j.biombioe.2014.10.016>.
- Negrutiu, I., R. Shillito, I. Potrykus, G. Biasini, and F. Sala. 1987. “Hybrid Genes in the Analysis of Transformation Conditions.” *Plant Molecular Biology* 8 (5): 363–73.
<https://doi.org/10.1007/BF00015814>.
- Nersissian, Aram M., and Eric L. Shipp. 2002. “Blue Copper-Binding Domains.” In , 271–340.

- [https://doi.org/10.1016/S0065-3233\(02\)60056-7](https://doi.org/10.1016/S0065-3233(02)60056-7).
- Newman, R. H., S. J. Hill, and P. J. Harris. 2013. "Wide-Angle X-Ray Scattering and Solid-State Nuclear Magnetic Resonance Data Combined to Test Models for Cellulose Microfibrils in Mung Bean Cell Walls." *Plant Physiology* 163 (4): 1558–67. <https://doi.org/10.1104/pp.113.228262>.
- Newman, Lisa J., Daniel E. Perazza, Lusanda Juda, and Malcolm M. Campbell. 2004. "Involvement of the R2R3-MYB, AtMYB61, in the Ectopic Lignification and Dark-Photomorphogenic Components of the Det3 Mutant Phenotype." *The Plant Journal* 37 (2): 239–50. <https://doi.org/10.1046/j.1365-313X.2003.01953.x>.
- Nguyen, Tuong-Van, Tran Thanh Thu, Martine Claeys, and Geert Angenon. 2007. "Agrobacterium - Mediated Transformation of Sorghum (Sorghum Bicolor (L.) Moench) Using an Improved in Vitro Regeneration System." *Plant Cell, Tissue and Organ Culture* 91 (2): 155–64. <https://doi.org/10.1007/s11240-007-9228-1>.
- Nguyen-Kim, H. (2015). "Recherche de la fonction de protéines riches en hydroxyproline dans les parois végétales." Laboratoire de Recherche en Sciences Végétales (LRSV), UMR 5546 UPS/CNRS, Université de Toulouse
- Nicholl, Desmond S. T. 2002. *An Introduction to Genetic Engineering*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139168205>.
- Nieuwland, Jeroen, Richard Feron, Bastiaan A.H. Huisman, Annalisa Fasolino, Cornelis W. Hilbers, Jan Derksen, and Celestina Mariani. 2005. "Lipid Transfer Proteins Enhance Cell Wall Extension in Tobacco." *The Plant Cell* 17 (7): 2009–19. <https://doi.org/10.1105/tpc.105.032094>.
- Nieves-Cordones, Manuel, Alain Chavanieu, Linda Jeanguenin, Carine Alcon, Wojciech Szponarski, Sebastien Estaran, Isabelle Chérel, Sabine Zimmermann, Hervé Sentenac, and Isabelle Gaillard. 2014. "Distinct Amino Acids in the C-Linker Domain of the Arabidopsis K + Channel KAT2 Determine Its Subcellular Localization and Activity at the Plasma Membrane." *Plant Physiology* 164 (3): 1415–29. <https://doi.org/10.1104/pp.113.229757>.
- Nishitani, Kazuhiko, and Donald J. Nevins. 1989. "Enzymic Analysis of Feruloylated Arabinoxylans (Feraxan) Derived from Zea Mays Cell Walls." *Plant Physiology* 91 (1): 242–48. <https://doi.org/10.1104/pp.91.1.242>.
- Nixon, B. Tracy, Katayoun Mansouri, Abhishek Singh, Juan Du, Jonathan K. Davis, Jung-Goo Lee, Erin Slabaugh, et al. 2016. "Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex." *Scientific Reports* 6 (1): 28696. <https://doi.org/10.1038/srep28696>.
- Norwood, Charles A., and Randall S. Currie. 1997. "Dryland Corn vs. Grain Sorghum in Western Kansas." *Jpa* 10 (1): 152. <https://doi.org/10.2134/jpa1997.0152>.
- Nuruzzaman, Mohammed, Ramaswamy Manimekalai, Akhter Most Sharoni, Kouji Satoh, Hiroaki Kondoh, Hisako Ooka, and Shoshi Kikuchi. 2010. "Genome-Wide Analysis of NAC Transcription Factor Family in Rice." *Gene* 465 (1–2): 30–44.

- <https://doi.org/10.1016/j.gene.2010.06.008>.
- Nuruzzaman, Mohammed, Akhter M. Sharoni, and Shoshi Kikuchi. 2013. "Roles of NAC Transcription Factors in the Regulation of Biotic and Abiotic Stress Responses in Plants." *Frontiers in Microbiology* 4 (SEP): 1–16. <https://doi.org/10.3389/fmicb.2013.00248>.
- O'Neill, Malcolm, Peter Albersheim, and Alan Darvill. 1990. "The Pectic Polysaccharides of Primary Cell Walls." In , 415–41. <https://doi.org/10.1016/B978-0-12-461012-5.50018-5>.
- Ochoa-Villarreal, Marisol, Emmanuel Aispuro-Hernandez, Irasema Vargas-Arispuro, and Miguel Ngel. 2012. "Plant Cell Wall Polymers: Function, Structure and Biological Activity of Their Derivatives." In *Polymerization*. InTech. <https://doi.org/10.5772/46094>.
- Odell, Joan T., Ferenc Nagy, and Nam-Hai Chua. 1985. "Identification of DNA Sequences Required for Activity of the Cauliflower Mosaic Virus 35S Promoter." *Nature* 313 (6005): 810–12. <https://doi.org/10.1038/313810a0>.
- Ohashi-Ito, Kyoko, Yoshihisa Oda, and Hiroo Fukuda. 2010. "Arabidopsis VASCULAR-RELATED NAC-DOMAIN6 Directly Regulates the Genes That Govern Programmed Cell Death and Secondary Wall Formation during Xylem Differentiation." *The Plant Cell* 22 (10): 3461–73. <https://doi.org/10.1105/tpc.110.075036>.
- Ohtani, Misato, Nobuyuki Nishikubo, Bo Xu, Masatoshi Yamaguchi, Nobutaka Mitsuda, Nadia Goué, Fusun Shi, Masaru Ohme-Takagi, and Taku Demura. 2011. "A NAC Domain Protein Family Contributing to the Regulation of Wood Formation in Poplar." *The Plant Journal* 67 (3): 499–512. <https://doi.org/10.1111/j.1365-313X.2011.04614.x>.
- Oldham, Michael C., Steve Horvath, and Daniel H. Geschwind. 2006. "Conservation and Evolution of Gene Coexpression Networks in Human and Chimpanzee Brains." *Proceedings of the National Academy of Sciences of the United States of America* 103 (47): 17973–78. <https://doi.org/10.1073/pnas.0605938103>.
- Ooka, Hisako, Hisako Ooka, Kouji Satoh, Kouji Satoh, Koji Doi, Koji Doi, Toshifumi Nagata, et al. 2003. "Comprehensive Analysis of NAC Family Genes In." *Analysis* 247: 239–47.
- Ou-Lee, T. M., R. Turgeon, and R. Wu. 1986. "Expression of a Foreign Gene Linked to Either a Plant-Virus or a Drosophila Promoter, after Electroporation of Protoplasts of Rice, Wheat, and Sorghum." *Proceedings of the National Academy of Sciences of the United States of America* 83 (18): 6815–19. <https://doi.org/10.1073/pnas.83.18.6815>.
- Pagant, Silvére, Adeline Bichet, Keiko Sugimoto, Olivier Lerouxel, Thierry Desprez, Maureen McCann, Patrice Lerouge, Samantha Vernhettes, and Herman Höfte. 2002. "Kobito1 Encodes a Novel Plasma Membrane Protein Necessary for Normal Synthesis of Cellulose during Cell Expansion in Arabidopsis." *Plant Cell* 14 (9): 2001–13. <https://doi.org/10.1105/tpc.002873>.
- Page, Mike T., Martin A.J. Parry, and Elizabete Carmo-Silva. 2019. "A High-throughput Transient Expression System for Rice." *Plant, Cell & Environment*, April, pce.13542. <https://doi.org/10.1111/pce.13542>.

- Palmer, Nathan A., Scott E. Sattler, Aaron J. Saathoff, Deanna Funnell, Jeffery F. Pedersen, and Gautam Sarath. 2008. "Genetic Background Impacts Soluble and Cell Wall-Bound Aromatics in Brown Midrib Mutants of Sorghum." *Planta* 229 (1): 115–27. <https://doi.org/10.1007/s00425-008-0814-1>.
- Panche, A. N., A. D. Diwan, and S. R. Chandra. 2016. "Flavonoids: An Overview." *Journal of Nutritional Science* 5 (December): e47. <https://doi.org/10.1017/jns.2016.41>.
- Paszkowski, Jerzy, Raymond D Shillito, Michael Saul, Vaclav Mandak, Thomas Hohn, and Barbara Hohn. 1984. "Direct Gene Transfer to Plants" 3 (12): 2717–22.
- Paterson, Andrew H., John E. Bowers, Rémy Bruggmann, Inna Dubchak, Jane Grimwood, Heidrun Gundlach, Georg Haberer, et al. 2009. "The Sorghum Bicolor Genome and the Diversification of Grasses." *Nature* 457 (7229): 551–56. <https://doi.org/10.1038/nature07723>.
- Pawar, Prashant Mohan-Anupama, Sanna Koutaniemi, Maija Tenkanen, and Ewa J. Mellerowicz. 2013. "Acetylation of Woody Lignocellulose: Significance and Regulation." *Frontiers in Plant Science* 4. <https://doi.org/10.3389/fpls.2013.00118>.
- Peaucelle, Alexis, Siobhan Braybrook, and Herman Höfte. 2012. "Cell Wall Mechanics and Growth Control in Plants: The Role of Pectins Revisited." *Frontiers in Plant Science* 3. <https://doi.org/10.3389/fpls.2012.00121>.
- Pechanova, Olga, Chuan-Yu Hsu, Joshua P Adams, Tibor Pechan, Lindsay Vandervelde, Jenny Drnevich, Sara Jawdy, et al. 2010. "Apoplast Proteome Reveals That Extracellular Matrix Contributes to Multistress Response in Poplar." *BMC Genomics* 11 (1): 674. <https://doi.org/10.1186/1471-2164-11-674>.
- Pedersen, J. F., J. J. Toy, D. L. Funnell, S. E. Sattler, A. L. Oliver, and R. A. Grant. 2008. "Registration of BN611, AN612, BN612, and RN613 Sorghum Genetic Stocks with Stacked -6 and -12 Genes." *Journal of Plant Registrations* 2 (3): 258. <https://doi.org/10.3198/jpr2008.01.0065crs>.
- Penning, Bryan W., Charles T. Hunter, Reuben Tayengwa, Andrea L. Eveland, Christopher K. Dugard, Anna T. Olek, Wilfred Vermerris, et al. 2009. "Genetic Resources for Maize Cell Wall Biology." *Plant Physiology* 151 (4): 1703–28. <https://doi.org/10.1104/pp.109.136804>.
- Pereira-Santana, Alejandro, Luis David Alcaraz, Enrique Castaño, Lenin Sanchez-Calderon, Felipe Sanchez-Teyer, and Luis Rodriguez-Zapata. 2015. "Comparative Genomics of NAC Transcriptional Factors in Angiosperms: Implications for the Adaptation and Diversification of Flowering Plants." *PLoS ONE* 10 (11). <https://doi.org/10.1371/journal.pone.0141866>.
- Perrier, L. (2017a). "Analyse et modélisation de la variabilité phénotypique du sorgho biomasse (Sorghum bicolor) pour l'exploration d'idéotypes dans un contexte de diversification des usages". Unité de Recherche AGAP, CIRAD, Montpellier
- Perrier, Lisa, Lauriane Rouan, Sylvie Jaffuel, Anne Clément-Vidal, Sandrine Roques, Armelle Soutiras, Christelle Baptiste, et al. 2017b. "Plasticity of Sorghum Stem Biomass Accumulation in

- Response to Water Deficit: A Multiscale Analysis from Internode Tissue to Plant Level.” *Frontiers in Plant Science* 8 (September). <https://doi.org/10.3389/fpls.2017.01516>.
- Perrin, R. M. 1999. “Xyloglucan Fucosyltransferase, an Enzyme Involved in Plant Cell Wall Biosynthesis.” *Science* 284 (5422): 1976–79. <https://doi.org/10.1126/science.284.5422.1976>.
- Persson, Staffan, Alexander Paredez, Andrew Carroll, Hildur Palsdottir, Monika Doblin, Patricia Poindexter, Natalie Khitrov, Manfred Auer, and Chris R. Somerville. 2007. “Genetic Evidence for Three Unique Components in Primary Cell-Wall Cellulose Synthase Complexes in Arabidopsis.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (39): 15566–71. <https://doi.org/10.1073/pnas.0706592104>.
- Persson, Staffan, Hairong Wei, Jennifer Milne, Grier P. Page, and Christopher R. Somerville. 2005. “Identification of Genes Required for Cellulose Synthesis by Regression Analysis of Public Microarray Data Sets.” *Proceedings of the National Academy of Sciences of the United States of America* 102 (24): 8633–38. <https://doi.org/10.1073/pnas.0503392102>.
- Pertea, Mihaela, Daehwan Kim, Geo M Pertea, Jeffrey T Leek, and Steven L Salzberg. 2016. “Transcript-Level Expression Analysis of RNA-Seq Experiments with HISAT, StringTie and Ballgown.” *Nature Protocols* 11 (August): 1650. <https://doi.org/10.1038/nprot.2016.095>.
- Pesch, Marina, and Martin Hülskamp. 2009. “One, Two, Three...models for Trichome Patterning in Arabidopsis?” *Current Opinion in Plant Biology* 12 (5): 587–92. <https://doi.org/10.1016/j.pbi.2009.07.015>.
- Petti, Carloalberto, Anne E. Harman-Ware, Mizuki Tateno, Rekha Kushwaha, Andrew Shearer, A. Bruce Downie, Mark Crocker, and Seth Debolt. 2013. “Sorghum Mutant RG Displays Antithetic Leaf Shoot Lignin Accumulation Resulting in Improved Stem Saccharification Properties.” *Biotechnology for Biofuels* 6 (1): 1. <https://doi.org/10.1186/1754-6834-6-146>.
- Petti, Carloalberto, Ko Hirano, Jozsef Stork, and Seth DeBolt. 2015. “Mapping of a Cellulose-Deficient Mutant Named Dwarf1-1 in Sorghum Bicolor to the Green Revolution Gene Gibberellin20-Oxidase Reveals a Positive Regulatory Association between Gibberellin and Cellulose Biosynthesis.” *Plant Physiology* 169 (1): 705–16. <https://doi.org/10.1104/pp.15.00928>.
- Plomion, C, G Leprovost, and A Stokes. 2001. “Wood Formation in Trees.” *Plant Physiology* 127 (4): 1513–23.
- Poehlman, John Milton. 1987. “Breeding Sorghum and Millet.” In *Breeding Field Crops*, 508–55. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-015-7271-2_19.
- Potrykus, Ingo, Michael W. Saul, Jirina Petruska, Jerzy Paszkowski, and Raymond D. Shillito. 1985. “Direct Gene Transfer to Cells of a Gramineous Monocot.” *MGG Molecular & General Genetics* 199 (2): 183–88. <https://doi.org/10.1007/BF00330257>.
- Preston, Jeremy, Janet Wheeler, Joshua Heazlewood, Song Feng Li, and Roger W. Parish. 2004. “AtMYB32 Is Required for Normal Pollen Development in Arabidopsis Thaliana.” *The Plant*

- Journal* 40 (6): 979–95. <https://doi.org/10.1111/j.1365-313X.2004.02280.x>.
- Proost, Sebastian, Agnieszka Krawczyk, and Marek Mutwil. 2017. “LSTrAP: Efficiently Combining RNA Sequencing Data into Co-Expression Networks.” *BMC Bioinformatics* 18 (1): 1–9. <https://doi.org/10.1186/s12859-017-1861-z>.
- Propheter, J. L., S. A. Staggenborg, X. Wu, and D. Wang. 2010. “Performance of Annual and Perennial Biofuel Crops: Yield during the First Two Years.” *Agronomy Journal* 102 (2): 806. <https://doi.org/10.2134/agronj2009.0301>.
- Pruneda-Paz, Jose L., Ghislain Breton, Dawn H. Nagel, S. Earl Kang, Katia Bonaldi, Colleen J. Doherty, Stephanie Ravelo, Mary Galli, Joseph R. Ecker, and Steve A. Kay. 2014. “A Genome-Scale Resource for the Functional Characterization of Arabidopsis Transcription Factors.” *Cell Reports* 8 (2): 622–32. <https://doi.org/10.1016/j.celrep.2014.06.033>.
- Pu, Xin-an, and R.N. Goodman. 1992. “Induction of Necrogenesis by *Agrobacterium Tumefaciens* on Grape Explants.” *Physiological and Molecular Plant Pathology* 41 (4): 241–54. [https://doi.org/10.1016/0885-5765\(92\)90024-P](https://doi.org/10.1016/0885-5765(92)90024-P).
- Puranik, Swati, Pranav Pankaj Sahu, Sambhu Nath Mandal, Venkata Suresh B., Swarup Kumar Parida, and Manoj Prasad. 2013. “Comprehensive Genome-Wide Survey, Genomic Constitution and Expression Profiling of the NAC Transcription Factor Family in Foxtail Millet (*Setaria Italica* L.).” *PLoS ONE* 8 (5). <https://doi.org/10.1371/journal.pone.0064594>.
- Puranik, Swati, Pranav Pankaj Sahu, Prem S. Srivastava, and Manoj Prasad. 2012. “NAC Proteins: Regulation and Role in Stress Tolerance.” *Trends in Plant Science* 17 (6): 369–81. <https://doi.org/10.1016/j.tplants.2012.02.004>.
- Rabinowicz, P D, E L Braun, A D Wolfe, B Bowen, and E Grotewold. 1999. “Maize R2R3 Myb Genes: Sequence Analysis Reveals Amplification in the Higher Plants.” *Genetics* 153 (1): 427–44. <http://www.ncbi.nlm.nih.gov/pubmed/10471724>.
- Raes, Jeroen, Antje Rohde, Jørgen Holst Christensen, Yves Van de Peer, and Wout Boerjan. 2003. “Genome-Wide Characterization of the Lignification Toolbox in Arabidopsis.” *Plant Physiology* 133 (3): 1051–71. <https://doi.org/10.1104/pp.103.026484>.
- Raghuwanshi, Anshu, and Robert G. Birch. 2010. “Genetic Transformation of Sweet Sorghum.” *Plant Cell Reports* 29 (9): 997–1005. <https://doi.org/10.1007/s00299-010-0885-x>.
- Rai, Krishan M., Sandi W. Thu, Vimal K. Balasubramanian, Christopher J. Cobos, Tesfaye Disasa, and Venugopal Mendu. 2016. “Identification, Characterization, and Expression Analysis of Cell Wall Related Genes in Sorghum Bicolor (L.) Moench, a Food, Fodder, and Biofuel Crop.” *Frontiers in Plant Science* 7 (August): 1–19. <https://doi.org/10.3389/fpls.2016.01287>.
- Rajashekar, C. B., and M. J. Burke. 1996. “Freezing Characteristics of Rigid Plant Tissues (Development of Cell Tension during Extracellular Freezing).” *Plant Physiology* 111 (2): 597–603. <https://doi.org/10.1104/pp.111.2.597>.
- Rajashekar, C. B., and A. Lafta. 1996. “Cell-Wall Changes and Cell Tension in Response to Cold

- Acclimation and Exogenous Absciscic Acid in Leaves and Cell Cultures.” *Plant Physiology* 111 (2): 605–12. <https://doi.org/10.1104/pp.111.2.605>.
- Rakoczy-Trojanowska, Monika. 2002. “Alternative Methods of Plant Transformation: A Short Review.” *Cellular & Molecular Biology Letters* 7 (3): 849–58.
- Ralph, John, Knut Lundquist, Gösta Brunow, Fachuang Lu, Hoon Kim, Paul F. Schatz, Jane M. Marita, et al. 2004. “Lignins: Natural Polymers from Oxidative Coupling of 4-Hydroxyphenyl-Propanoids.” *Phytochemistry Reviews* 3 (1–2): 29–60. <https://doi.org/10.1023/B:PHYT.0000047809.65444.a4>.
- Ranjan, Alok, Neha Pandey, Deepika Lakhwani, Neeraj Kumar Dubey, Uday V Pathre, and Samir V Sawant. 2012. “Comparative Transcriptomic Analysis of Roots of Contrasting Gossypium Herbaceum Genotypes Revealing Adaptation to Drought.” *BMC Genomics* 13 (1): 680. <https://doi.org/10.1186/1471-2164-13-680>.
- Rao, Xiaolan, Xin Chen, Hui Shen, Qin Ma, Guifen Li, Yuhong Tang, Maria Pena, et al. 2019. “Gene Regulatory Networks for Lignin Biosynthesis in Switchgrass (*Panicum Virgatum*).” *Plant Biotechnology Journal* 17 (3): 580–93. <https://doi.org/10.1111/pbi.13000>.
- Rao, Xiaolan, and Richard A. Dixon. 2018. “Current Models for Transcriptional Regulation of Secondary Cell Wall Biosynthesis in Grasses.” *Frontiers in Plant Science* 9 (April): 1–11. <https://doi.org/10.3389/fpls.2018.00399>.
- Rao, Xiaolan, Hui Shen, Sivakumar Pattathil, Michael G. Hahn, Ivana Gelineo-Albersheim, Debra Mohnen, Yunqiao Pu, et al. 2017. “Dynamic Changes in Transcriptome and Cell Wall Composition Underlying Brassinosteroid-Mediated Lignification of Switchgrass Suspension Cells.” *Biotechnology for Biofuels* 10 (1): 266. <https://doi.org/10.1186/s13068-017-0954-2>.
- Ratanasut, Kumrop, Weerawan Rod-In, and Kawee Sujipuli. 2017. “In Planta Agrobacterium - Mediated Transformation of Rice.” *Rice Science* 24 (3): 181–86. <https://doi.org/10.1016/j.rsci.2016.11.001>.
- Read, S. M., and A. Bacic. 2002. “PLANT BIOLOGY: Prime Time for Cellulose.” *Science* 295 (5552): 59–60. <https://doi.org/10.1126/science.1068155>.
- Rennie, Emilie A, and Henrik Vibe Scheller. 2014. “Xylan Biosynthesis.” *Current Opinion in Biotechnology* 26 (April): 100–107. <https://doi.org/10.1016/j.copbio.2013.11.013>.
- Repellin, Anne, Monica Båga, Prem P. Jauhar, and Ravindra N. Chibbar. 2001. “Genetic Enrichment of Cereal Crops via Alien Gene Transfer: New Challenges.” *Plant Cell, Tissue and Organ Culture* 64 (2–3): 159–83. <https://doi.org/10.1023/A:1010633510352>.
- Richie, W.R, and G.G McBee. 1991. “Structural Components in Sorghum Stem Biomass.” *Bioresource Technology* 38 (1): 15–22. [https://doi.org/10.1016/0960-8524\(91\)90216-7](https://doi.org/10.1016/0960-8524(91)90216-7).
- Ridley, Brent L, Malcolm A O'Neill, and Debra Mohnen. 2001. “Pectins: Structure, Biosynthesis, and Oligogalacturonide-Related Signaling.” *Phytochemistry* 57 (6): 929–67. [https://doi.org/10.1016/S0031-9422\(01\)00113-3](https://doi.org/10.1016/S0031-9422(01)00113-3).

- Ritchie, Matthew E., Belinda Phipson, Di Wu, Yifang Hu, Charity W. Law, Wei Shi, and Gordon K. Smyth. 2015. "Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies." *Nucleic Acids Research* 43 (7): e47. <https://doi.org/10.1093/nar/gkv007>.
- Robert, Stéphanie, Adeline Bichet, Olivier Grandjean, Daniel Kierzkowski, Béatrice Satiat-Jeunemaître, Sandra Pelletier, Marie-Theres Hauser, Herman Höfte, and Samantha Vernhettes. 2005. "An Arabidopsis Endo-1,4- β -D-Glucanase Involved in Cellulose Synthesis Undergoes Regulated Intracellular Cycling." *The Plant Cell* 17 (12): 3378–89. <https://doi.org/10.1105/tpc.105.036228>.
- Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. "EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data." *Bioinformatics* 26 (1): 139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
- Rocateli, A.C., R.L. Raper, K.S. Balkcom, F.J. Arriaga, and D.I. Bransby. 2012. "Biomass Sorghum Production and Components under Different Irrigation/Tillage Systems for the Southeastern U.S." *Industrial Crops and Products* 36 (1): 589–98. <https://doi.org/10.1016/j.indcrop.2011.11.007>.
- Romano, Julia M., Christian Dubos, Michael B. Prouse, Olivia Wilkins, Henry Hong, Mervin Poole, Kyu-Young Kang, et al. 2012. "At MYB61, an R2R3-MYB Transcription Factor, Functions as a Pleiotropic Regulator via a Small Gene Network." *New Phytologist* 195 (4): 774–86. <https://doi.org/10.1111/j.1469-8137.2012.04201.x>.
- Rooney, William L., Jürg Blumenthal, Brent Bean, and John E. Mullet. 2007. "Designing Sorghum as a Dedicated Bioenergy Feedstock." *Biofuels, Bioproducts and Biorefining* 1 (2): 147–57. <https://doi.org/10.1002/bbb.15>.
- Rose, Jocelyn K. C., Janet Braam, Stephen C. Fry, and Kazuhiko Nishitani. 2002. "The XTH Family of Enzymes Involved in Xyloglucan Endotransglucosylation and Endohydrolysis: Current Perspectives and a New Unifying Nomenclature." *Plant and Cell Physiology* 43 (12): 1421–35. <https://doi.org/10.1093/pcp/pcf171>.
- Rosenow, D.T., J.E. Quisenberry, C.W. Wendt, and L.E. Clark. 1983. "Drought Tolerant Sorghum and Cotton Germplasm." *Agricultural Water Management* 7 (1–3): 207–22. [https://doi.org/10.1016/0378-3774\(83\)90084-7](https://doi.org/10.1016/0378-3774(83)90084-7).
- Rosinski, James A., and William R. Atchley. 1998. "Molecular Evolution of the Myb Family of Transcription Factors: Evidence for Polyphyletic Origin." *Journal of Molecular Evolution* 46 (1): 74–83. <https://doi.org/10.1007/PL00006285>.
- Rubio, Vicente, Francisco Linhares, Roberto Solano, Ana C. Martín, Joaquín Iglesias, Antonio Leyva, and Javier Paz-Ares. 2001. "A Conserved MYB Transcription Factor Involved in Phosphate Starvation Signaling Both in Vascular Plants and in Unicellular Algae." *Genes and Development* 15 (16): 2122–33. <https://doi.org/10.1101/gad.204401>.
- Ruel, Katia, Yoshiharu Nishiyama, and Jean-Paul Joseleau. 2012. "Crystalline and Amorphous

- Cellulose in the Secondary Walls of Arabidopsis.” *Plant Science* 193–194 (September): 48–61. <https://doi.org/10.1016/j.plantsci.2012.05.008>.
- Ruprecht, Colin, Marek Mutwil, Friederike Saxe, Michaela Eder, Zoran Nikoloski, and Staffan Persson. 2011. “Large-Scale Co-Expression Approach to Dissect Secondary Cell Wall Formation Across Plant Species.” *Frontiers in Plant Science* 2. <https://doi.org/10.3389/fpls.2011.00023>.
- Saballos, Ana. 2008. “Development and Utilization of Sorghum as a Bioenergy Crop.” In *Genetic Improvement of Bioenergy Crops*, 211–48. New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-70805-8_8.
- Saballos, Ana, Gebisa Ejeta, Emiliano Sanchez, ChulHee Kang, and Wilfred Vermerris. 2009. “A Genomewide Analysis of the Cinnamyl Alcohol Dehydrogenase Family in Sorghum [Sorghum Bicolor (L.) Moench] Identifies SbCAD2 as the Brown Midrib6 Gene.” *Genetics* 181 (2): 783–95. <https://doi.org/10.1534/genetics.108.098996>.
- Saballos, Ana, Scott E. Sattler, Emiliano Sanchez, Timothy P. Foster, Zhanguo Xin, Chulhee Kang, Jeffrey F. Pedersen, and Wilfred Vermerris. 2012. “Brown Midrib2 (Bmr2) Encodes the Major 4-Coumarate: Coenzyme A Ligase Involved in Lignin Biosynthesis in Sorghum (Sorghum Bicolor (L.) Moench).” *Plant Journal* 70 (5): 818–30. <https://doi.org/10.1111/j.1365-3113.2012.04933.x>.
- Sablowski, Robert W.M, and Elliot M Meyerowitz. 1998. “A Homolog of NO APICAL MERISTEM Is an Immediate Target of the Floral Homeotic Genes APETALA3/PISTILLATA.” *Cell* 92 (1): 93–103. [https://doi.org/10.1016/S0092-8674\(00\)80902-2](https://doi.org/10.1016/S0092-8674(00)80902-2).
- Sáez-Plaza, Purificación, María José Navas, Sławomir Wybraniec, Tadeusz Michałowski, and Agustín García Asuero. 2013. “An Overview of the Kjeldahl Method of Nitrogen Determination. Part II. Sample Preparation, Working Scale, Instrumental Finish, and Quality Control.” *Critical Reviews in Analytical Chemistry* 43 (4): 224–72. <https://doi.org/10.1080/10408347.2012.751787>.
- Saha, Badal C. 2003. “Hemicellulose Bioconversion.” *Journal of Industrial Microbiology and Biotechnology* 30 (5): 279–91. <https://doi.org/10.1007/s10295-003-0049-x>.
- Saha, Gopal, Jong In Park, Nasar Uddin Ahmed, Md Abdul Kayum, Kwon Kyoo Kang, and Ill Sup Nou. 2016. “Characterization and Expression Profiling of MYB Transcription Factors against Stresses and during Male Organ Development in Chinese Cabbage (Brassica Rapa Ssp. Pekinensis).” *Plant Physiology and Biochemistry* 104: 200–215. <https://doi.org/10.1016/j.plaphy.2016.03.021>.
- Sairam, R. V., N. Seetharama, P. S. Devi, A. Verma, U. R. Murthy, and I. Potrykus. 1999. “Culture and Regeneration of Mesophyll-Derived Protoplasts of Sorghum [Sorghum Bicolor (L.) Moench].” *Plant Cell Reports* 18 (12): 972–77. <https://doi.org/10.1007/s002990050693>.
- Samuels, Lacey, Ljerka Kunst, and Reinhard Jetter. 2008. “Sealing Plant Surfaces: Cuticular Wax Formation by Epidermal Cells.” *Annual Review of Plant Biology* 59 (1): 683–707. <https://doi.org/10.1146/annurev.arplant.59.103006.093219>.

- Sanagala, Raghavendrarao, Anil Kumar Moola, and Ranjitha Kumari Bollipo Diana. 2017. "A Review on Advanced Methods in Plant Gene Targeting." *Journal of Genetic Engineering and Biotechnology* 15 (2): 317–21. <https://doi.org/10.1016/j.jgeb.2017.07.004>.
- Sánchez-Rodríguez, Clara, Stefan Bauer, Kian Hématy, Friederike Saxe, Ana Belén Ibáñez, Vera Vordermaier, Cornelia Konlechner, et al. 2012. "CHITINASE-LIKE1/POM-POM1 and Its Homolog CTL2 Are Glucan-Interacting Proteins Important for Cellulose Biosynthesis in Arabidopsis." *The Plant Cell* 24 (2): 589–607. <https://doi.org/10.1105/tpc.111.094672>.
- Sanford, John C., Theodore M. Klein, Edward D. Wolf, And Nelson Allen. 1987. "Delivery Of Substances Into Cells And Tissues Using A Particle Bombardment Process." *Particulate Science and Technology* 5 (1): 27–37. <https://doi.org/10.1080/02726358708904533>.
- Sanjari, Sepideh, Reza Shirzadian-Khorramabad, Zahra Sadat Shobbar, and Maryam Shahbazi. 2019. "Systematic Analysis of NAC Transcription Factors' Gene Family and Identification of Post-Flowering Drought Stress Responsive Members in Sorghum." *Plant Cell Reports* 38 (3): 361–76. <https://doi.org/10.1007/s00299-019-02371-8>.
- Sant, Rajnesh R. Prasad. 2011. "Development of a Transformation System for Sorghum (Sorghum Bicolor L.)." Center for Tropical Crops and Biocommodities, Queensland University of Technology
- Sarah, Gautier, Felix Homa, Stéphanie Pointet, Sandy Contreras, François Sabot, Benoit Nabholz, Sylvain Santoni, et al. 2017. "A Large Set of 26 New Reference Transcriptomes Dedicated to Comparative Population Genomics in Crops and Wild Relatives." *Molecular Ecology Resources* 17 (3): 565–80. <https://doi.org/10.1111/1755-0998.12587>.
- Sasaki, Takuji, and International Rice Genome Sequencing Project. 2005. "The Map-Based Sequence of the Rice Genome." *Nature* 436 (7052): 793–800. <https://doi.org/10.1038/nature03895>.
- Sasidharan, Rashmi, Laurentius A C J Voesenek, and Ronald Pierik. 2011. "Cell Wall Modifying Proteins Mediate Plant Acclimatization to Biotic and Abiotic Stresses." *Critical Reviews in Plant Sciences* 30 (6): 548–62. <https://doi.org/10.1080/07352689.2011.615706>.
- Sato, Kanna, Ryu Suzuki, Nobuyuki Nishikubo, Sachi Takenouchi, Sachiko Ito, Yoshimi Nakano, Satoshi Nakaba, et al. 2010. "Isolation of a Novel Cell Wall Architecture Mutant of Rice with Defective Arabidopsis COBL4 Ortholog BC1 Required for Regulated Deposition of Secondary Cell Wall Components." *Planta* 232 (1): 257–70. <https://doi.org/10.1007/s00425-010-1171-4>.
- Sato, S., T. Clemente, and I. Dweikat. 2004. "Identification of an Elite Sorghum Genotype with High In Vitro Performance Capacity." *In Vitro Cellular & Developmental Biology - Plant* 40 (1): 57–60. <https://doi.org/10.1079/IVP2003475>.
- Sattler, Scott E., and Deanna L. Funnell-Harris. 2013. "Modifying Lignin to Improve Bioenergy Feedstocks: Strengthening the Barrier against Pathogens?†." *Frontiers in Plant Science* 4. <https://doi.org/10.3389/fpls.2013.00070>.
- Sattler, Scott E., Ana Saballos, Zhanguo Xin, Deanna L. Funnell-Harris, Wilfred Vermerris, and

- Jeffrey F. Pedersen. 2014. “ Characterization of Novel Sorghum Brown Midrib Mutants from an EMS-Mutagenized Population .” *G3: Genes/Genomes/Genetics* 4 (11): 2115–24. <https://doi.org/10.1534/g3.114.014001>.
- Saxena, Inder M., And R. Malcolm Brown. 2005. “Cellulose Biosynthesis: Current Views and Evolving Concepts.” *Annals of Botany* 96 (1): 9–21. <https://doi.org/10.1093/aob/mci155>.
- Schaeffer, Pierre, and Rollin D. Hotchkiss. 1978. “Fusion of Bacterial Protoplasts.” *Methods in Cell Biology* 20 (C): 149–58. [https://doi.org/10.1016/S0091-679X\(08\)62017-8](https://doi.org/10.1016/S0091-679X(08)62017-8).
- Scheller, Henrik Vibe, and Peter Ulvskov. 2010. “Hemicelluloses.” *Annual Review of Plant Biology* 61 (1): 263–89. <https://doi.org/10.1146/annurev-arplant-042809-112315>.
- Schindelman, G. 2001. “COBRA Encodes a Putative GPI-Anchored Protein, Which Is Polarly Localized and Necessary for Oriented Cell Expansion in Arabidopsis.” *Genes & Development* 15 (9): 1115–27. <https://doi.org/10.1101/gad.879101>.
- Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler. 2002. “TREE-PUZZLE: Maximum Likelihood Phylogenetic Analysis Using Quartets and Parallel Computing.” *Bioinformatics* 18 (3): 502–4. <https://doi.org/10.1093/bioinformatics/18.3.502>.
- Schmitz, Aaron J., Kevin Begcy, Gautam Sarath, and Harkamal Walia. 2015. “Rice Ovate Family Protein 2 (OFP2) Alters Hormonal Homeostasis and Vasculature Development.” *Plant Science* 241 (December): 177–88. <https://doi.org/10.1016/j.plantsci.2015.10.011>.
- Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, et al. 2009. “The B73 Maize Genome: Complexity, Diversity, and Dynamics.” *Science* 326 (5956): 1112–15. <https://doi.org/10.1126/science.1178534>.
- Schopfer, P. 2006. “Biomechanics of Plant Growth.” *American Journal of Botany* 93 (10): 1415–25. <https://doi.org/10.3732/ajb.93.10.1415>.
- Scully, Erin D., Tammy Gries, Gautam Sarath, Nathan A. Palmer, Lisa Baird, Michelle J. Serapiglia, Bruce S. Dien, et al. 2016a. “Overexpression of SbMyb60 Impacts Phenylpropanoid Biosynthesis and Alters Secondary Cell Wall Composition in Sorghum Bicolor.” *Plant Journal* 85 (3): 378–95. <https://doi.org/10.1111/tpj.13112>.
- Scully, Erin D., Tammy Gries, Deanna L. Funnell-Harris, Zhanguo Xin, Frank A. Kovacs, Wilfred Vermerris, and Scott E. Sattler. 2016b. “Characterization of Novel Brown Midrib 6 Mutations Affecting Lignin Biosynthesis in Sorghum.” *Journal of Integrative Plant Biology* 58 (2): 136–49. <https://doi.org/10.1111/jipb.12375>.
- Scully, Erin D., Tammy Gries, Nathan A. Palmer, Gautam Sarath, Deanna L. Funnell-Harris, Lisa Baird, Paul Twigg, Javier Seravalli, Thomas E. Clemente, and Scott E. Sattler. 2017. “Overexpression of SbMyb60 in Sorghum Bicolor Impacts Both Primary and Secondary Metabolism.” *New Phytologist* 63. <https://doi.org/10.1111/nph.14815>.
- Seifert, Georg J., and Claudia Blaukopf. 2010. “Irritable Walls: The Plant Extracellular Matrix and Signaling.” *Plant Physiology* 153 (2): 467–78. <https://doi.org/10.1104/pp.110.153940>.

- Seifert, Georg J., and Keith Roberts. 2007. "The Biology of Arabinogalactan Proteins." *Annual Review of Plant Biology* 58 (1): 137–61. <https://doi.org/10.1146/annurev.arplant.58.032806.103801>.
- Shakoor, Nadia, Ramesh Nair, Oswald Crasta, Geoffrey Morris, Alex Feltus, and Stephen Kresovich. 2014. "A Sorghum Bicolor Expression Atlas Reveals Dynamic Genotype-Specific Expression Profiles for Vegetative Tissues of Grain, Sweet and Bioenergy Sorghums." *BMC Plant Biology* 14 (1): 35. <https://doi.org/10.1186/1471-2229-14-35>.
- Sheen, J. 2001. "Signal Transduction in Maize and Arabidopsis Mesophyll Protoplasts." *Plant Physiology* 127 (4): 1466–75.
- Shen, Hui, Xianzhi He, Charleson R. Poovaiah, Wegi A. Wuddineh, Junying Ma, David G. J. Mann, Huanzhong Wang, et al. 2012. "Functional Characterization of the Switchgrass (*Panicum Virgatum*) R2R3-MYB Transcription Factor PvMYB4 for Improvement of Lignocellulosic Feedstocks." *New Phytologist* 193 (1): 121–36. <https://doi.org/10.1111/j.1469-8137.2011.03922.x>.
- Shen, Hui, Yanbin Yin, Fang Chen, Ying Xu, and Richard A. Dixon. 2009. "A Bioinformatic Analysis of NAC Genes for Plant Cell Wall Development in Relation to Lignocellulosic Bioenergy Production." *BioEnergy Research* 2 (4): 217–32. <https://doi.org/10.1007/s12155-009-9047-9>.
- Shimada, Setsuko, Yuko Makita, Tomoko Kuriyama-Kondou, Mika Kawashima, Yoshiki Mochizuki, Hideki Hirakawa, Shusei Sato, Tetsuro Toyoda, and Minami Matsui. 2015. "Functional and Expression Analyses of Transcripts Based on Full-Length cDNAs of Sorghum Bicolor." *DNA Research* 22 (6): 485–93. <https://doi.org/10.1093/dnares/dsv030>.
- Shimamoto, Ko, Rie Terada, Takeshi Izawa, and Hideya Fujimoto. 1989. "Fertile Transgenic Rice Plants Regenerated from Transformed Protoplasts." *Nature* 338 (6212): 274–76. <https://doi.org/10.1038/338274a0>.
- Shiriga, Kaliyugam, Rinku Sharma, Krishan Kumar, Shiv Kumar Yadav, Firoz Hossain, and Nepolean Thirunavukkarasu. 2014. "Genome-Wide Identification and Expression Pattern of Drought-Responsive Members of the NAC Family in Maize." *Meta Gene* 2 (1): 407–17. <https://doi.org/10.1016/j.mgene.2014.05.001>.
- Shiringani, Amukelani L., and Wolfgang Friedt. 2011. "QTL for Fibre-Related Traits in Grain × Sweet Sorghum as a Tool for the Enhancement of Sorghum as a Biomass Crop." *Theoretical and Applied Genetics* 123 (6): 999–1011. <https://doi.org/10.1007/s00122-011-1642-4>.
- Shiu, S.-H., and A. B. Blecker. 2001. "Receptor-like Kinases from Arabidopsis Form a Monophyletic Gene Family Related to Animal Receptor Kinases." *Proceedings of the National Academy of Sciences* 98 (19): 10763–68. <https://doi.org/10.1073/pnas.181141598>.
- Shrawat, Ashok Kumar, and Horst Lörz. 2006. "Agrobacterium-Mediated Transformation of Cereals: A Promising Approach Crossing Barriers." *Plant Biotechnology Journal* 4 (6): 575–603. <https://doi.org/10.1111/j.1467-7652.2006.00209.x>.

- Sibout, Richard, Sebastian Proost, Bjoern Oest Hansen, Neha Vaid, Federico M. Giorgi, Severine Ho-Yue-Kuang, Frédéric Legée, et al. 2017a. “Expression Atlas and Comparative Coexpression Network Analyses Reveal Important Genes Involved in the Formation of Lignified Cell Wall in *Brachypodium Distachyon*.” *New Phytologist* 215 (3): 1009–25. <https://doi.org/10.1111/nph.14635>.
- Silva, Laura S. da, Janet Taylor, and John R. N. Taylor. 2011. “Transgenic Sorghum with Altered Kafirin Synthesis: Kafirin Solubility, Polymerization, and Protein Digestion.” *Journal of Agricultural and Food Chemistry* 59 (17): 9265–70. <https://doi.org/10.1021/jf201878p>.
- Simon, Marissa, Myeong Min Lee, Yan Lin, Lindsey Gish, and John Schiefelbein. 2007. “Distinct and Overlapping Roles of Single-Repeat MYB Genes in Root Epidermal Patterning.” *Developmental Biology* 311 (2): 566–78. <https://doi.org/10.1016/j.ydbio.2007.09.001>.
- Sinclair, Rosalie, Michel Ruiz Rosquete, and Georgia Drakakaki. 2018. “Post-Golgi Trafficking and Transport of Cell Wall Components.” *Frontiers in Plant Science* 871 (December): 1–11. <https://doi.org/10.3389/fpls.2018.01784>.
- Sircar, Sanchari, Nita Parekh, and Gaurav Sablok. 2018. “Protocol for Coexpression Network Construction and Stress-Responsive Expression Analysis in *Brachypodium*.” In *Methods in Molecular Biology (Clifton, N.J.)*, edited by Gaurav Sablok, Hikmet Budak, and Peter J. Ralph, 1667:203–21. *Methods in Molecular Biology*. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4939-7278-4_16.
- Skadhauge, Birgitte, Karl Kristian Thomsen, and Diter Wettstein. 1997. “The Role of the Barley Testa Layer and Its Flavonoid Content in Resistance to *Fusarium* Infections.” *Hereditas* 126 (2): 147–60. <https://doi.org/10.1111/j.1601-5223.1997.00147.x>.
- Skirycz, Aleksandra, and Dirk Inzé. 2010. “More from Less: Plant Growth under Limited Water.” *Current Opinion in Biotechnology* 21 (2): 197–203. <https://doi.org/10.1016/j.copbio.2010.03.002>.
- Slater A, Scott N, Fowler M. 2003. “Plant Biotechnology: The Genetic Manipulation of Plants.” *Annals of Botany* 94 (4): 646–646. <https://doi.org/10.1093/aob/mch186>.
- Slewinski, T. L. 2012. “Non-Structural Carbohydrate Partitioning in Grass Stems: A Target to Increase Yield Stability, Stress Tolerance, and Biofuel Production.” *Journal of Experimental Botany* 63 (13): 4647–70. <https://doi.org/10.1093/jxb/ers124>.
- Smallwood, Maggie, and Dianna J. Bowles. 2002. “Plants in a Cold Climate.” Edited by D. J. Bowles, P. J. Lillford, D. A. Rees, and I. A. Shanks. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 357 (1423): 831–47. <https://doi.org/10.1098/rstb.2002.1073>.
- Smita, Shuchi, Amit Katiyar, Viswanathan Chinnusamy, Dev M. Pandey, and Kailash C. Bansal. 2015. “Transcriptional Regulatory Network Analysis of MYB Transcription Factor Family Genes in Rice.” *Frontiers in Plant Science* 6 (December).

- <https://doi.org/10.3389/fpls.2015.01157>.
- Smith, G. A., M. O. Bagby, R. T. Lewellan, D. L. Doney, P. H. Moore, F. J. Hills, L. G. Campbell, G. J. Hogaboam, G. E. Coe, and K. Freeman. 1987. "Evaluation of Sweet Sorghum for Fermentable Sugar Production Potential1." *Crop Science* 27 (4): 788. <https://doi.org/10.2135/cropsci1987.0011183x002700040037x>.
- Smith, Rebecca A., Mathias Schuetz, Steven D. Karlen, David Bird, Naohito Tokunaga, Yasushi Sato, Shawn D. Mansfield, John Ralph, and A. Lacey Samuels. 2017. "Defining the Diverse Cell Populations Contributing to Lignification in Arabidopsis Stems." *Plant Physiology* 174 (2): 1028–36. <https://doi.org/10.1104/pp.17.00434>.
- Smith, Roberta H., and Elizabeth E. Hood. 1995. "Agrobacterium Tumefaciens Transformation of Monocotyledons." *Crop Science* 35 (2): 301. <https://doi.org/10.2135/cropsci1995.0011183X003500020001x>.
- Soest, P.J. Van, J.B. Robertson, and B.A. Lewis. 1991. "Methods for Dietary Fiber, Neutral Detergent Fiber, and Nonstarch Polysaccharides in Relation to Animal Nutrition." *Journal of Dairy Science* 74 (10): 3583–97. [https://doi.org/10.3168/jds.S0022-0302\(91\)78551-2](https://doi.org/10.3168/jds.S0022-0302(91)78551-2).
- Solecka, Danuta, Alain-M Boudet, and Alina Kacperska. 1999. "Phenylpropanoid and Anthocyanin Changes in Low-Temperature Treated Winter Oilseed Rape Leaves." *Plant Physiology and Biochemistry* 37 (6): 491–96. [https://doi.org/10.1016/S0981-9428\(99\)80054-0](https://doi.org/10.1016/S0981-9428(99)80054-0).
- Solecka, Danuta, Jacek Żebrowski, and Alina Kacperska. 2008. "Are Pectins Involved in Cold Acclimation and De-Acclimation of Winter Oil-Seed Rape Plants?" *Annals of Botany* 101 (4): 521–30. <https://doi.org/10.1093/aob/mcm329>.
- Soler, Marçal, Eduardo Leal Oliveira Camargo, Victor Carocha, Hua Cassan-Wang, Hélène San Clemente, Bruno Savelli, Charles A Hefer, Jorge A Pinto Paiva, Alexander A Myburg, and Jacqueline Grima-Pettenati. 2015. "The Eucalyptus Grandis R2R3-MYB Transcription Factor Family: Evidence for Woody Growth-Related Evolution and Function." *New Phytologist* 206 (4): 1364–77. <https://doi.org/10.1111/nph.13039>.
- Sonbol, Fathi-Mohamed, Silvia Fornalé, Montserrat Capellades, Antonio Encina, Sonia Touriño, Josep-Lluís Torres, Pere Rovira, et al. 2009. "The Maize ZmMYB42 Represses the Phenylpropanoid Pathway and Affects the Cell Wall Structure, Composition and Degradability in Arabidopsis Thaliana." *Plant Molecular Biology* 70 (3): 283–96. <https://doi.org/10.1007/s11103-009-9473-2>.
- Sood, P., A. Bhattacharya, and A. Sood. 2011. "Problems and Possibilities of Monocot Transformation." *Biologia Plantarum* 55 (1): 1–15. <https://doi.org/10.1007/s10535-011-0001-2>.
- Souer, Erik, Adèle van Houwelingen, Daisy Kloos, Jos Mol, and Ronald Koes. 1996. "The No Apical Meristem Gene of Petunia Is Required for Pattern Formation in Embryos and Flowers and Is Expressed at Meristem and Primordia Boundaries." *Cell* 85 (2): 159–70. [https://doi.org/10.1016/S0092-8674\(00\)81093-4](https://doi.org/10.1016/S0092-8674(00)81093-4).

- Souza, Wagner R. de, Polyana K. Martins, Jackie Freeman, Till K. Pellny, Louise V. Michaelson, Bruno L. Sampaio, Felipe Vinecky, et al. 2018. "Suppression of a Single BAHD Gene in *Setaria Viridis* Causes Large, Stable Decreases in Cell Wall Feruloylation and Increases Biomass Digestibility." *New Phytologist* 218 (1): 81–93. <https://doi.org/10.1111/nph.14970>.
- Stefanowska, M. 1999. "Low Temperature Affects Pattern of Leaf Growth and Structure of Cell Walls in Winter Oilseed Rape (*Brassica Napus* L., Var. *Oleifera* L.)." *Annals of Botany* 84 (3): 313–19. <https://doi.org/10.1006/anbo.1999.0924>.
- Streatfield, Stephen J, Maria E Magallanes-Lundback, Katherine K Beifuss, Christopher A Brooks, Robin L Harkey, Robert T Love, Jeff Bray, John A Howard, Joseph M Jilka, and Elizabeth E Hood. 2004. "Analysis of the Maize Polyubiquitin-1 Promoter Heat Shock Elements and Generation of Promoter Variants with Modified Expression Characteristics." *Transgenic Research* 13 (4): 299–312.
- Sutter, Valerie, Rudy Vanderhaeghen, Sofie Tilleman, Freya Lammertyn, Isabelle Vanhoutte, Mansour Karimi, Dirk Inzé, Alain Goossens, and Pierre Hilson. 2005. "Exploration of Jasmonate Signalling via Automated and Standardized Transient Expression Assays in Tobacco Cells." *The Plant Journal* 44 (6): 1065–76. <https://doi.org/10.1111/j.1365-313X.2005.02586.x>.
- Swigonova, Z. 2004. "Close Split of Sorghum and Maize Genome Progenitors." *Genome Research* 14 (10a): 1916–23. <https://doi.org/10.1101/gr.2332504>.
- Szyjanowicz, Pio M. J., Iain McKinnon, Neil G. Taylor, John Gardiner, Mike C. Jarvis, and Simon R. Turner. 2004. "The Irregular Xylem 2 Mutant Is an Allele of Korrigan That Affects the Secondary Cell Wall of *Arabidopsis Thaliana*." *The Plant Journal* 37 (5): 730–40. <https://doi.org/10.1111/j.1365-313X.2003.02000.x>.
- Tadesse, Yohannes, László Sági, Rony Swennen, and Michel Jacobs. 2003. "Optimisation of Transformation Conditions and Production of Transgenic Sorghum (*Sorghum Bicolor*) via Microparticle Bombardment." *Plant Cell, Tissue and Organ Culture* 75 (1): 1–18. <https://doi.org/10.1023/A:1024664817800>.
- Tai, Yuling, Chun Liu, Shuwei Yu, Hua Yang, Jiameng Sun, Chunxiao Guo, Bei Huang, et al. 2018. "Gene Co-Expression Network Analysis Reveals Coordinated Regulation of Three Characteristic Secondary Biosynthetic Pathways in Tea Plant (*Camellia Sinensis*)." *BMC Genomics* 19 (1): 1–13. <https://doi.org/10.1186/s12864-018-4999-9>.
- Taiz, L. and Zeiger, E. (1998) "Plant Physiology." 2nd Edition, Sinauer Associates Publishers, Sunderland, Massachusetts.
- Takahagi, Kotaro, Komaki Inoue, and Keiichi Mochida. 2018. "Gene Co-Expression Network Analysis Suggests the Existence of Transcriptional Modules Containing a High Proportion of Transcriptionally Differentiated Homoeologs in Hexaploid Wheat." *Frontiers in Plant Science* 9 (August): 1–10. <https://doi.org/10.3389/fpls.2018.01163>.
- Tan, Mingpu, Dan Cheng, Yuening Yang, Guoqiang Zhang, Mengjie Qin, Jun Chen, Yahua Chen, and

- Mingyi Jiang. 2017. “Co-Expression Network Analysis of the Transcriptomes of Rice Roots Exposed to Various Cadmium Stresses Reveals Universal Cadmium-Responsive Genes.” *BMC Plant Biology* 17 (1): 1–18. <https://doi.org/10.1186/s12870-017-1143-y>.
- Tan, Tian Tian, Hitoshi Endo, Ryosuke Sano, Tetsuya Kurata, Masatoshi Yamaguchi, Misato Ohtani, and Taku Demura. 2018. “Transcription Factors VND1-VND3 Contribute to Cotyledon Xylem Vessel Formation.” *Plant Physiology* 176 (1): 773–89. <https://doi.org/10.1104/pp.17.00461>.
- Tanaka, Katsuyuki, Kazumasa Murata, Muneo Yamazaki, Katsura Onosato, Akio Miyao, and Hirohiko Hirochika. 2003. “Three Distinct Rice Cellulose Synthase Catalytic Subunit Genes Required for Cellulose Synthesis in the Secondary Wall.” *Plant Physiology* 133 (1): 73–83. <https://doi.org/10.1104/pp.103.022442>.
- Taylor, N. G., R. M. Howells, A. K. Huttly, K. Vickers, and S. R. Turner. 2003. “Interactions among Three Distinct CesaA Proteins Essential for Cellulose Synthesis.” *Proceedings of the National Academy of Sciences* 100 (3): 1450–55. <https://doi.org/10.1073/pnas.0337628100>.
- Tenhaken, Raimund. 2015. “Cell Wall Remodeling under Abiotic Stress.” *Frontiers in Plant Science* 5 (JAN): 1–9. <https://doi.org/10.3389/fpls.2014.00771>.
- Terashima, N., H. G. Jung, D. R. Buxton, R. D. Hatfield, J. Ralph, K. Fukushima, L-F. He, and K. Takabe. 1993. “Comprehensive Model of the Lignified Plant Cell Wall.” In . <https://doi.org/10.2134/1993.foragecellwall.c10>.
- Tetreault, Hannah M., Erin D. Scully, Tammy Gries, Nathan A. Palmer, Deanna L. Funnell-Harris, Lisa Baird, Javier Seravalli, et al. 2018a. “Overexpression of the Sorghum Bicolor SbCCoAOMT Alters Cell Wall Associated Hydroxycinnamoyl Groups.” *PLoS ONE* 13 (10): 1–25. <https://doi.org/10.1371/journal.pone.0204153>.
- Thera, K. (2017). "Analyse des déterminants génétiques contrôlant la production et la composition de la tige chez le sorgho (*Sorghum Bicolor* [L.] Moench). Intégration des approches bi- et multiparentales. Unité de Recherche AGAP, CIRAD, Montpellier
- Thomma, Bart, Bruno Cammue, and Karin Thevissen. 2002. “Plant Defensins.” *Planta* 216 (2): 193–202. <https://doi.org/10.1007/s00425-002-0902-6>.
- Thompson, Dawn, Aviv Regev, and Sushmita Roy. 2015. “Comparative Analysis of Gene Regulatory Networks: From Network Reconstruction to Evolution.” *Annual Review of Cell and Developmental Biology* 31 (1): 399–428. <https://doi.org/10.1146/annurev-cellbio-100913-012908>.
- Thornton, Philip K., Anthony Whitbread, Tobias Baedeker, Jill Cairns, Lieven Claessens, Walter Baethgen, Christian Bunn, et al. 2018. “A Framework for Priority-Setting in Climate Smart Agriculture Research.” *Agricultural Systems* 167 (November): 161–75. <https://doi.org/10.1016/j.agsy.2018.09.009>.
- Tian, Tian, Qi You, Liwei Zhang, Xin Yi, Hengyu Yan, and Wenying Xu. 2016. “Original Article SorghumFDB: Sorghum Functional Genomics Database with Multidimensional Network

- Analysis,” 1–16. <https://doi.org/10.1093/database/baw099>.
- Tingay, Sonia, David McElroy, Roger Kalla, Sarah Fieg, Mingbo Wang, Sarah Thornton, and Richard Brettell. 1997. “Agrobacterium Tumefaciens-Mediated Barley Transformation.” *The Plant Journal* 11 (6): 1369–76. <https://doi.org/10.1046/j.1365-313X.1997.11061369.x>.
- Tobimatsu, Yuki, and Mathias Schuetz. 2019. “Lignin Polymerization: How Do Plants Manage the Chemistry so Well?” *Current Opinion in Biotechnology* 56 (April): 75–81. <https://doi.org/10.1016/j.copbio.2018.10.001>.
- Todaka, D., K. Nakashima, K. Maruyama, S. Kidokoro, Y. Osakabe, Y. Ito, S. Matsukura, et al. 2012. “Rice Phytochrome-Interacting Factor-like Protein OsPIL1 Functions as a Key Regulator of Internode Elongation and Induces a Morphological Response to Drought Stress.” *Proceedings of the National Academy of Sciences* 109 (39): 15947–52. <https://doi.org/10.1073/pnas.1207324109>.
- Townsley, Brad T., Neelima R. Sinha, and Julie Kang. 2013. “KNOX1 Genes Regulate Lignin Deposition and Composition in Monocots and Dicots.” *Frontiers in Plant Science* 4. <https://doi.org/10.3389/fpls.2013.00121>.
- Tran, Lam-Son Phan, Kazuo Nakashima, Yoh Sakuma, Sean D. Simpson, Yasunari Fujita, Kyonoshin Maruyama, Miki Fujita, Motoaki Seki, Kazuo Shinozaki, and Kazuko Yamaguchi-Shinozaki. 2004. “Isolation and Functional Analysis of Arabidopsis Stress-Inducible NAC Transcription Factors That Bind to a Drought-Responsive Cis -Element in the Early Responsive to Dehydration Stress 1 Promoter.” *The Plant Cell* 16 (9): 2481–98. <https://doi.org/10.1105/tpc.104.022699>.
- Treutter, D. 2005. “Significance of Flavonoids in Plant Resistance and Enhancement of Their Biosynthesis.” *Plant Biology* 7 (6): 581–91. <https://doi.org/10.1055/s-2005-873009>.
- Tripoli, Elisa, Maurizio La Guardia, Santo Giammanco, Danila Di Majo, and Marco Giammanco. 2007. “Citrus Flavonoids: Molecular Structure, Biological Activity and Nutritional Properties: A Review.” *Food Chemistry* 104 (2): 466–79. <https://doi.org/10.1016/j.foodchem.2006.11.054>.
- Tronchet, Maurice, Claudine Balagué, Thomas Kroj, Lise Jouanin, and Dominique Roby. 2010. “Cinnamyl Alcohol Dehydrogenases-C and D, Key Enzymes in Lignin Biosynthesis, Play an Essential Role in Disease Resistance in Arabidopsis.” *Molecular Plant Pathology* 11 (1): 83–92. <https://doi.org/10.1111/j.1364-3703.2009.00578.x>.
- Trouche, G., D. Bastianelli, T. V.Cao Hamadou, J. Chantereau, J. F. Rami, and D. Pot. 2014a. “Exploring the Variability of a Photoperiod-Insensitive Sorghum Genetic Panel for Stem Composition and Related Traits in Temperate Environments.” *Field Crops Research* 166: 72–81. <https://doi.org/10.1016/j.fcr.2014.06.008>.
- Trouche, G., D. Bastianelli, T. V Cao Hamadou, J. Chantereau, J. F. Rami, and D. Pot. 2014b. “Exploring the Variability of a Photoperiod-Insensitive Sorghum Genetic Panel for Stem Composition and Related Traits in Temperate Environments.” *Field Crops Research* 166: 72–81. <https://doi.org/10.1016/j.fcr.2014.06.008>.

- Tzfira, Tzvi, and Vitaly Citovsky. 2006. "Agrobacterium-Mediated Genetic Transformation of Plants: Biology and Biotechnology." *Current Opinion in Biotechnology* 17 (2): 147–54. <https://doi.org/10.1016/j.copbio.2006.01.009>.
- Uchimiya, Hirofumi, Takaomi Fushimi, Hiroyuki Hashimoto, Hiroshi Harada, Kunihiko Syono, and Yukihiro Sugawara. 1986. "Expression of a Foreign Gene in Callus Derived from DNA-Treated Protoplasts of Rice (*Oryza Sativa* L.)." *MGG Molecular & General Genetics* 204 (2): 204–7. <https://doi.org/10.1007/BF00425499>.
- Underwood, William. 2012. "The Plant Cell Wall: A Dynamic Barrier against Pathogen Invasion." *Frontiers in Plant Science* 3 (MAY): 1–6. <https://doi.org/10.3389/fpls.2012.00085>.
- Urbanowicz, Breeanna R., Maria J. Peña, Heather A. Moniz, Kelley W. Moremen, and William S. York. 2014. "Two Arabidopsis Proteins Synthesize Acetylated Xylan in Vitro." *The Plant Journal* 80 (2): 197–206. <https://doi.org/10.1111/tpj.12643>.
- Vain, Thomas, Elizabeth Faris Crowell, Hélène Timpano, Eric Biot, Thierry Desprez, Nasim Mansoori, Luisa M. Trindade, et al. 2014. "The Cellulase KORRIGAN Is Part of the Cellulose Synthase Complex." *Plant Physiology* 165 (4): 1521–32. <https://doi.org/10.1104/pp.114.241216>.
- Valdivia, Elene R., María Teresa Herrera, Cristina Gianzo, Javier Fidalgo, Gloria Revilla, Ignacio Zarra, and Javier Sampedro. 2013. "Regulation of Secondary Wall Synthesis and Cell Death by NAC Transcription Factors in the Monocot Brachypodium Distachyon." *Journal of Experimental Botany* 64 (5): 1333–43. <https://doi.org/10.1093/jxb/ers394>.
- Vandenbrink, Joshua P., Valorie Goff, Huizhe Jin, Wenqian Kong, Andrew H. Paterson, and F. Alex Feltus. 2013. "Identification of Bioconversion Quantitative Trait Loci in the Interspecific Cross *Sorghum Bicolor* ?? *Sorghum Propinquum*." *Theoretical and Applied Genetics* 126 (9): 2367–80. <https://doi.org/10.1007/s00122-013-2141-6>.
- Vanholme, Ruben, Igor Cesarino, Katarzyna Rataj, Yuguo Xiao, Lisa Sundin, Geert Goeminne, Hoon Kim, et al. 2013. "Caffeoyl Shikimate Esterase (CSE) Is an Enzyme in the Lignin Biosynthetic Pathway in Arabidopsis." *Science* 341 (6150): 1103–6. <https://doi.org/10.1126/science.1241602>.
- Vanholme, Ruben, Véronique Storme, Bartel Vanholme, Lisa Sundin, Jørgen Holst Christensen, Geert Goeminne, Claire Halpin, Antje Rohde, Kris Morreel, and Wout Boerjana. 2012. "A Systems Biology View of Responses to Lignin Biosynthesis Perturbations in Arabidopsis." *Plant Cell* 24 (9): 3506–29. <https://doi.org/10.1105/tpc.112.102574>.
- Vega-Sánchez, Miguel E., Dominique Loqué, Jeemeng Lao, Michela Catena, Yves Verhertbruggen, Thomas Herter, Fan Yang, et al. 2015. "Engineering Temporal Accumulation of a Low Recalcitrance Polysaccharide Leads to Increased C6 Sugar Content in Plant Cell Walls." *Plant Biotechnology Journal* 13 (7): 903–14. <https://doi.org/10.1111/pbi.12326>.
- Vélez-Bermúdez, Isabel-Cristina, Jorge E. Salazar-Henao, Silvia Fornalé, Irene López-Vidriero, José-Manuel Franco-Zorrilla, Erich Grotewold, John Gray, et al. 2015. "A MYB/ZML Complex Regulates Wound-Induced Lignin Genes in Maize." *The Plant Cell* 27 (11): 3245–59.

- <https://doi.org/10.1105/tpc.15.00545>.
- Vermerris, Wilfred. 2011. “Survey of Genomics Approaches to Improve Bioenergy Traits in Maize, Sorghum and SugarcaneFree Access.” *Journal of Integrative Plant Biology* 53 (2): 105–19. <https://doi.org/10.1111/j.1744-7909.2010.01020.x>.
- Vermerris, Wilfred, Ana Saballos, Gebisa Ejeta, Nathan S. Mosier, Michael R. Ladisch, and Nicholas C. Carpita. 2007. “Molecular Breeding to Enhance Ethanol Production from Corn and Sorghum Stover.” *Crop Science* 47 (December): S-142-S-153. <https://doi.org/10.2135/cropsci2007.04.0013IPBS>.
- Veronese, Paola, Maria T. Ruiz, Maria A. Coca, Agustin Hernandez-Lopez, Hyeseung Lee, José I. Ibeas, Barbara Damsz, et al. 2003. “In Defense against Pathogens. Both Plant Sentinels and Foot Soldiers Need to Know the Enemy: Fig. 1.” *Plant Physiology* 131 (4): 1580–90. <https://doi.org/10.1104/pp.102.013417>.
- Vo, Loan T.T., Jordi Girones, Calypso Beloli, Lucie Chupin, Erika Di Giuseppe, Anne Clément Vidal, Armelle Soutiras, et al. 2017. “Processing and Properties of Sorghum Stem Fragment-Polyethylene Composites.” *Industrial Crops and Products* 107 (November): 386–98. <https://doi.org/10.1016/j.indcrop.2017.05.047>.
- Voesenek, L. A. C. J., T. D. Colmer, R. Pierik, F. F. Millenaar, and A. J. M. Peeters. 2006. “How Plants Cope with Complete Submergence.” *New Phytologist* 170 (2): 213–26. <https://doi.org/10.1111/j.1469-8137.2006.01692.x>.
- Vogel, John. 2008. “Unique Aspects of the Grass Cell Wall.” *Current Opinion in Plant Biology* 11 (3): 301–7. <https://doi.org/10.1016/j.pbi.2008.03.002>.
- Vreeburg, Robert A.M., Joris J. Benschop, Anton J.M. Peeters, Timothy D. Colmer, Ankie H.M. Ammerlaan, Marten Staal, Theo M. Elzenga, et al. 2005. “Ethylene Regulates Fast Apoplastic Acidification and Expansin A Transcription during Submergence-Induced Petiole Elongation in *Rumex Palustris*.” *The Plant Journal* 43 (4): 597–610. <https://doi.org/10.1111/j.1365-313X.2005.02477.x>.
- Wakabayashi, Kazuyuki, Kouichi Soga, and Takayuki Hoson. 2012. “Phenylalanine Ammonia-Lyase and Cell Wall Peroxidase Are Cooperatively Involved in the Extensive Formation of Ferulate Network in Cell Walls of Developing Rice Shoots.” *Journal of Plant Physiology* 169 (3): 262–67. <https://doi.org/10.1016/j.jplph.2011.10.002>.
- Waldron, C., E. B. Murphy, J. L. Roberts, G. D. Gustafson, S. L. Armour, and S. K. Malcolm. 1985. “Resistance to Hygromycin B.” *Plant Molecular Biology* 5 (2): 103–8. <https://doi.org/10.1007/BF00020092>.
- Wan, Y., and P. G. Lemaux. 1994. “Generation of Large Numbers of Independently Transformed Fertile Barley Plants.” *Plant Physiology* 104 (1): 37–48. <https://doi.org/10.1104/pp.104.1.37>.
- Wang, H., U. Avci, J. Nakashima, M. G. Hahn, F. Chen, and R. A. Dixon. 2010. “Mutation of WRKY Transcription Factors Initiates Pith Secondary Wall Formation and Increases Stem Biomass in

- Dicotyledonous Plants.” *Proceedings of the National Academy of Sciences* 107 (51): 22338–43. <https://doi.org/10.1073/pnas.1016436107>.
- Wang, Hai, Melvin Tucker, and Yun Ji. 2013. “Recent Development in Chemical Depolymerization of Lignin: A Review.” *Journal of Applied Chemistry* 2013: 1–9. <https://doi.org/10.1155/2013/838645>.
- Wang, Haifeng, Marie La Russa, and Lei S. Qi. 2016. “CRISPR/Cas9 in Genome Editing and Beyond.” *Annual Review of Biochemistry* 85 (1): 227–64. <https://doi.org/10.1146/annurev-biochem-060815-014607>.
- Wang, Hengsheng, Longjiang Gu, Xingen Zhang, Mingli Liu, Haiyang Jiang, Ronghao Cai, Yang Zhao, and Beijiu Cheng. 2018. “Global Transcriptome and Weighted Gene Co-Expression Network Analyses Reveal Hybrid-Specific Modules and Candidate Genes Related to Plant Height Development in Maize.” *Plant Molecular Biology* 98 (3): 187–203. <https://doi.org/10.1007/s11103-018-0763-4>.
- Wang, Huan-Zhong, and Richard A. Dixon. 2012. “On–Off Switches for Secondary Cell Wall Biosynthesis.” *Molecular Plant* 5 (2): 297–303. <https://doi.org/10.1093/mp/ssr098>.
- Wang, Huanzhong, Utku Avci, Jin Nakashima, Michael G Hahn, Fang Chen, and Richard A Dixon. 2010. “Mutation of WRKY Transcription Factors Initiates Pith Secondary Wall Formation and Increases Stem Biomass in Dicotyledonous Plants.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (51): 22338–43. <https://doi.org/10.1073/pnas.1016436107>.
- Wang, Ming L., Chengsong Zhu, Noelle A. Barkley, Zhenbang Chen, John E. Erpelding, Seth C. Murray, Mitchell R. Tuinstra, Tesfaye Tesso, Gary A. Pederson, and Jianming Yu. 2009. “Genetic Diversity and Population Structure Analysis of Accessions in the US Historic Sweet Sorghum Collection.” *Theoretical and Applied Genetics* 120 (1): 13–23. <https://doi.org/10.1007/s00122-009-1155-6>.
- Wang, Yi-Hong, Aniruddha Acharya, A Millie Burrell, Robert R Klein, Patricia E Klein, and Karl H Hasenstein. 2013. “Mapping and Candidate Genes Associated with Saccharification Yield in Sorghum.” *Genome* 56 (11): 659–65. <https://doi.org/10.1139/gen-2013-0134>.
- Wang, Yi Hong, Aniruddha Acharya, A. Millie Burrell, Robert R. Klein, Patricia E. Klein, and Karl H. Hasenstein. 2013. “Mapping and Candidate Genes Associated with Saccharification Yield in Sorghum.” *Genome* 56 (11): 659–65. <https://doi.org/10.1139/gen-2013-0134>.
- Wang, Zhi-Ping, Hui-Li Xing, Li Dong, Hai-Yan Zhang, Chun-Yan Han, Xue-Chen Wang, and Qi-Jun Chen. 2015. “Egg Cell-Specific Promoter-Controlled CRISPR/Cas9 Efficiently Generates Homozygous Mutants for Multiple Target Genes in Arabidopsis in a Single Generation.” *Genome Biology* 16 (1): 144. <https://doi.org/10.1186/s13059-015-0715-0>.
- Waters, Mark T, and Jane A Langdale. 2009. “The Making of a Chloroplast.” *The EMBO Journal* 28 (19): 2861–73. <https://doi.org/10.1038/emboj.2009.264>.

- Wei, Z. M., and Z. H. Xu. 1993. "Regeneration of Plants from Protoplasts of Sorghum (*Sorghum Vulgare*)." In , 123–31. https://doi.org/10.1007/978-3-642-78037-0_10.
- Wilkins, Olivia, Hardeep Nahal, Justin Foong, Nicholas J. Provart, and Malcolm M. Campbell. 2009. "Expansion and Diversification of the Populus R2R3-MYB Family of Transcription Factors ." *Plant Physiology* 149 (2): 981–93. <https://doi.org/10.1104/pp.108.132795>.
- Willemsen, Viola, Marion Bauch, Tom Bennett, Ana Campilho, Harald Wolkenfelt, Jian Xu, Jim Haseloff, and Ben Scheres. 2008. "The NAC Domain Transcription Factors FEZ and SOMBRERO Control the Orientation of Cell Division Plane in Arabidopsis Root Stem Cells." *Developmental Cell* 15 (6): 913–22. <https://doi.org/10.1016/j.devcel.2008.09.019>.
- Wilson, J. R., and R. D. Hatfield. 1997. "Structural and Chemical Changes of Cell Wall Types during Stem Development: Consequences for Fibre Degradation by Rumen Microflora." *Australian Journal of Agricultural Research* 48 (2): 165. <https://doi.org/10.1071/A96051>.
- Winchell, Frank, Michael Brass, Andrea Manzo, Alemseged Beldados, Valentina Perna, Charlene Murphy, Chris Stevens, and Dorian Q. Fuller. 2018. "On the Origins and Dissemination of Domesticated Sorghum and Pearl Millet across Africa and into India: A View from the Butana Group of the Far Eastern Sahel." *African Archaeological Review* 35 (4): 483–505. <https://doi.org/10.1007/s10437-018-9314-2>.
- Wittkopp, P. J. 2007. "Variable Gene Expression in Eukaryotes: A Network Perspective." *Journal of Experimental Biology* 210 (9): 1567–75. <https://doi.org/10.1242/jeb.002592>.
- Wohlleben, W., W. Arnold, I. Broer, D. Hillemann, E. Strauch, and A. Punier. 1988. "Nucleotide Sequence of the Phosphinothricin N-Acetyltransferase Gene from *Streptomyces Viridochromogenes* Tü494 and Its Expression in *Nicotiana Tabacum*." *Gene* 70 (1): 25–37. [https://doi.org/10.1016/0378-1119\(88\)90101-1](https://doi.org/10.1016/0378-1119(88)90101-1).
- Wolf, S, K Hematy, and H Hofte. 2012. "Growth Control and Cell Wall Signaling in Plants." *Annu Rev Plant Biol* 63: 381–407. <https://doi.org/10.1146/annurev-arplant-042811-105449>.
- Wu, Ai-Min, Christophe Rihouey, Martial Seveno, Emma Hörnblad, Sunil Kumar Singh, Toshiro Matsunaga, Tadashi Ishii, Patrice Lerouge, and Alan Marchant. 2009. "The Arabidopsis IRX10 and IRX10-LIKE Glycosyltransferases Are Critical for Glucuronoxylan Biosynthesis during Secondary Cell Wall Formation." *The Plant Journal* 57 (4): 718–31. <https://doi.org/10.1111/j.1365-313X.2008.03724.x>.
- Wu, Emily, Brian Lenderts, Kimberly Glassman, Maya Berezowska-Kaniewska, Heather Christensen, Tracy Asmus, Shifu Zhen, Uyen Chu, Myeong Je Cho, and Zuo Yu Zhao. 2014. "Optimized Agrobacterium-Mediated Sorghum Transformation Protocol and Molecular Data of Transgenic Sorghum Plants." *In Vitro Cellular and Developmental Biology - Plant* 50 (1): 9–18. <https://doi.org/10.1007/s11627-013-9583-z>.
- Wu, Y., R. E. Sharp, D. M. Durachko, and D. J. Cosgrove. 1996. "Growth Maintenance of the Maize Primary Root at Low Water Potentials Involves Increases in Cell-Wall Extension Properties,

- Expansin Activity, and Wall Susceptibility to Expansins.” *Plant Physiology* 111 (3): 765–72. <https://doi.org/10.1104/pp.111.3.765>.
- Wuddineh, Wegi A., Mitra Mazarei, Ji-Yi Zhang, Geoffrey B. Turner, Robert W. Sykes, Stephen R. Decker, Mark F. Davis, Michael K. Udvardi, and C. Neal Stewart. 2016. “Identification and Overexpression of a Knotted1-Like Transcription Factor in Switchgrass (*Panicum Virgatum* L.) for Lignocellulosic Feedstock Improvement.” *Frontiers in Plant Science* 7 (April). <https://doi.org/10.3389/fpls.2016.00520>.
- Xia, Jingnu, Yunjun Zhao, Payne Burks, Markus Pauly, and Patrick J. Brown. 2018. “A Sorghum NAC Gene Is Associated with Variation in Biomass Properties and Yield Potential.” *Plant Direct* 2 (7): 1–11. <https://doi.org/10.1002/pld3.70>.
- Xiao, Wenhan, Yue Yang, and Jingjuan Yu. 2018. “ZmNST3 and ZmNST4 Are Master Switches for Secondary Wall Deposition in Maize (*Zea Mays* L.).” *Plant Science* 266: 83–94. <https://doi.org/10.1016/j.plantsci.2017.03.012>.
- Xie, Q., G Frugis, D Colgan, and N H Chua. 2000. “Arabidopsis NAC1 Transduces Auxin Signal Downstream of TIR1 to Promote Lateral Root Development.” *Genes & Development* 14 (23): 3024–36. <https://doi.org/10.1101/gad.852200>.
- Xin, Z., and J. Browse. 2000. “Cold Comfort Farm: The Acclimation of Plants to Freezing Temperatures.” *Plant, Cell & Environment* 23 (9): 893–902. <https://doi.org/10.1046/j.1365-3040.2000.00611.x>.
- Xin, Zhanguo, Ming Li Wang, Noelle A Barkley, Gloria Burow, Cleve Franks, Gary Pederson, and John Burke. 2008. “Applying Genotyping (TILLING) and Phenotyping Analyses to Elucidate Gene Function in a Chemically Induced Sorghum Mutant Population.” *BMC Plant Biology* 8 (1): 103. <https://doi.org/10.1186/1471-2229-8-103>.
- Xiong, Yuqing, Tieyan Liu, Chaoguang Tian, Shouhong Sun, Jiayang Li, and Mingsheng Chen. 2005. “Transcription Factors in Rice: A Genome-Wide Comparative Analysis between Monocots and Eudicots.” *Plant Molecular Biology* 59 (1): 191–203. <https://doi.org/10.1007/s11103-005-6503-6>.
- Xu, Hua, Anming Ding, Sihui Chen, Prince Marowa, Dian Wang, Min Chen, Ruibo Hu, et al. 2018. “Genome-Wide Analysis of Sorghum GT47 Family Reveals Functional Divergences of MUR3-Like Genes.” *Frontiers in Plant Science* 9 (December). <https://doi.org/10.3389/fpls.2018.01773>.
- Yamada, Tomoyoshi, Katsushi Kuroda, Yutaka Jitsuyama, Daisuke Takezawa, Keita Arakawa, and Seizo Fujikawa. 2002. “Roles of the Plasma Membrane and the Cell Wall in the Responses of Plant Cells to Freezing.” *Planta* 215 (5): 770–78. <https://doi.org/10.1007/s00425-002-0814-5>.
- Yamaguchi, Masatoshi, Minoru Kubo, Hiroo Fukuda, and Taku Demura. 2008. “VASCULAR-RELATED NAC-DOMAIN7 Is Involved in the Differentiation of All Types of Xylem Vessels in Arabidopsis Roots and Shoots.” *The Plant Journal* 55 (4): 652–64. <https://doi.org/10.1111/j.1365-313X.2008.03533.x>.

- Yamaguchi, Masatoshi, Nobutaka Mitsuda, Misato Ohtani, Masaru Ohme-Takagi, Ko Kato, and Taku Demura. 2011. "VASCULAR-RELATED NAC-DOMAIN 7 Directly Regulates the Expression of a Broad Range of Genes for Xylem Vessel Formation." *Plant Journal* 66 (4): 579–90. <https://doi.org/10.1111/j.1365-313X.2011.04514.x>.
- Yamaguchi, Masatoshi, Misato Ohtani, Nobutaka Mitsuda, Minoru Kubo, Masaru Ohme-Takagi, Hiroo Fukuda, and Taku Demura. 2010. "VND-INTERACTING2, a NAC Domain Transcription Factor, Negatively Regulates Xylem Vessel Formation in Arabidopsis." *The Plant Cell* 22 (4): 1249–63. <https://doi.org/10.1105/tpc.108.064048>.
- Yamaguchi, Miki, Haruka Fujimoto, Ko Hirano, Satoko Araki-Nakamura, Kozue Ohmae-Shinohara, Akihiro Fujii, Masako Tsunashima, et al. 2016. "Sorghum Dw1, an Agronomically Important Gene for Lodging Resistance, Encodes a Novel Protein Involved in Cell Proliferation." *Scientific Reports* 6 (February): 28366. <https://doi.org/10.1038/srep28366>.
- Yang, Chunhua, Dayong Li, Xue Liu, Chengjun Ji, Lili Hao, Xianfeng Zhao, Xiaobing Li, Caiyan Chen, Zhukuan Cheng, and Lihuang Zhu. 2014. "OsMYB103L, an R2R3-MYB Transcription Factor, Influences Leaf Rolling and Mechanical Strength in Rice (*Oryza Sativa* L.)." *BMC Plant Biology* 14 (1): 158. <https://doi.org/10.1186/1471-2229-14-158>.
- Yang, Li, Xin Zhao, Fan Yang, Di Fan, Yuanzhong Jiang, and Keming Luo. 2016. "PtrWRKY19, a Novel WRKY Transcription Factor, Contributes to the Regulation of Pith Secondary Wall Formation in *Populus Trichocarpa*." *Scientific Reports* 6 (1): 18643. <https://doi.org/10.1038/srep18643>.
- Yang, Ziheng, and Bruce Rannala. 2012. "Molecular Phylogenetics: Principles and Practice." *Nature Reviews Genetics* 13 (5): 303–14. <https://doi.org/10.1038/nrg3186>.
- Yanhui, Chen, Yang Xiaoyuan, He Kun, Liu Meihua, Li Jigang, Gao Zhaofeng, Lin Zhiqiang, et al. 2006. "The MYB Transcription Factor Superfamily of Arabidopsis: Expression Analysis and Phylogenetic Comparison with the Rice MYB Family." *Plant Molecular Biology* 60 (1): 107–24. <https://doi.org/10.1007/s11103-005-2910-y>.
- Yazawa, Takayuki, Hiroyuki Kawahigashi, Takashi Matsumoto, and Hiroshi Mizuno. 2013. "Simultaneous Transcriptome Analysis of Sorghum and Bipolaris Sorghicola by Using RNA-Seq in Combination with De Novo Transcriptome Assembly." Edited by Dawn Arnold. *PLoS ONE* 8 (4): e62460. <https://doi.org/10.1371/journal.pone.0062460>.
- Ye, Yafeng, Binmei Liu, Meng Zhao, Kun Wu, Weimin Cheng, Xiangbin Chen, Qian Liu, Zan Liu, Xiangdong Fu, and Yuejin Wu. 2015. "CEF1/OsMYB103L Is Involved in GA-Mediated Regulation of Secondary Wall Biosynthesis in Rice." *Plant Molecular Biology* 89 (4–5): 385–401. <https://doi.org/10.1007/s11103-015-0376-0>.
- Yeats, Trevor H., and Jocelyn K.C. Rose. 2008. "The Biochemistry and Biology of Extracellular Plant Lipid-Transfer Proteins (LTPs)." *Protein Science* 17 (2): 191–98. <https://doi.org/10.1110/ps.073300108>.

- Yepes, Luz Marcela, and Herb S. Aldwinckle. 1994. "Micropropagation of Thirteen Malus Cultivars and Rootstocks, and Effect of Antibiotics on Proliferation." *Plant Growth Regulation* 15 (1): 55–67. <https://doi.org/10.1007/BF00024677>.
- Yin, Yanbin, Jinling Huang, and Ying Xu. 2009. "The Cellulose Synthase Superfamily in Fully Sequenced Plants and Algae." *BMC Plant Biology* 9 (1): 99. <https://doi.org/10.1186/1471-2229-9-99>.
- Yoo, Sang-Dong, Young-Hee Cho, and Jen Sheen. 2007. "Arabidopsis Mesophyll Protoplasts: A Versatile Cell System for Transient Gene Expression Analysis." *Nature Protocols* 2 (7): 1565–72. <https://doi.org/10.1038/nprot.2007.199>.
- Yoshida, Kouki, Shingo Sakamoto, Tetsushi Kawai, Yoshinori Kobayashi, Kazuhito Sato, Yasunori Ichinose, Katsuro Yaoi, et al. 2013. "Engineering the Oryza Sativa Cell Wall with Rice NAC Transcription Factors Regulating Secondary Wall Formation." *Frontiers in Plant Science* 4 (October): 1–12. <https://doi.org/10.3389/fpls.2013.00383>.
- You, Jun, Lihua Zhang, Bo Song, Xiaoquan Qi, and Zhulong Chan. 2015. "Systematic Analysis and Identification of Stress-Responsive Genes of the NAC Gene Family in Brachypodium Distachyon." *PLoS ONE* 10 (3): 1–20. <https://doi.org/10.1371/journal.pone.0122027>.
- Yu, Hua, Bingke Jiao, and Chengzhi Liang. 2017. "High-Quality Rice RNA-Seq-Based Co-Expression Network for Predicting Gene Function and Regulation." *BioRxiv*, 138040. <https://doi.org/10.1101/138040>.
- Zale, Janice M., S. Agarwal, S. Loar, and C. M. Steber. 2009. "Evidence for Stable Transformation of Wheat by Floral Dip in Agrobacterium Tumefaciens." *Plant Cell Reports* 28 (6): 903–13. <https://doi.org/10.1007/s00299-009-0696-0>.
- Zang, Aiping, Xiaojie Xu, Steven Neill, and Weiming Cai. 2010. "Overexpression of OsRAN2 in Rice and Arabidopsis Renders Transgenic Plants Hypersensitive to Salinity and Osmotic Stress." *Journal of Experimental Botany* 61 (3): 777–89. <https://doi.org/10.1093/jxb/erp341>.
- Zhang, Baolong, Tongmin Zhao, Wengui Yu, Beiqing Kuang, Yuan Yao, Tingli Liu, Xiaoyang Chen, Wenhua Zhang, and Ai-Min Wu. 2014. "Functional Conservation of the Glycosyltransferase Gene GT47A in the Monocot Rice." *Journal of Plant Research* 127 (3): 423–32. <https://doi.org/10.1007/s10265-014-0631-5>.
- Zhang, Deshui, Maria Hrmova, Chun Hua Wan, Chunfa Wu, Jace Balzen, Wendy Cai, Jing Wang, et al. 2004. "Members of a New Group of Chitinase-like Genes Are Expressed Preferentially in Cotton Cells with Secondary Walls." *Plant Molecular Biology* 54 (3): 353–72. <https://doi.org/10.1023/B:PLAN.0000036369.55253.dd>.
- Zhang, Jin, Meng Xie, Gerald A. Tuskan, Wellington Muchero, and Jin-Gui Chen. 2018. "Recent Advances in the Transcriptional Regulation of Secondary Cell Wall Biosynthesis in the Woody Plants." *Frontiers in Plant Science* 9 (October). <https://doi.org/10.3389/fpls.2018.01535>.
- Zhang, Li-Min, Chuan-Yuan Leng, Hong Luo, Xiao-Yuan Wu, Zhi-Quan Liu, Yu-Miao Zhang, Hong

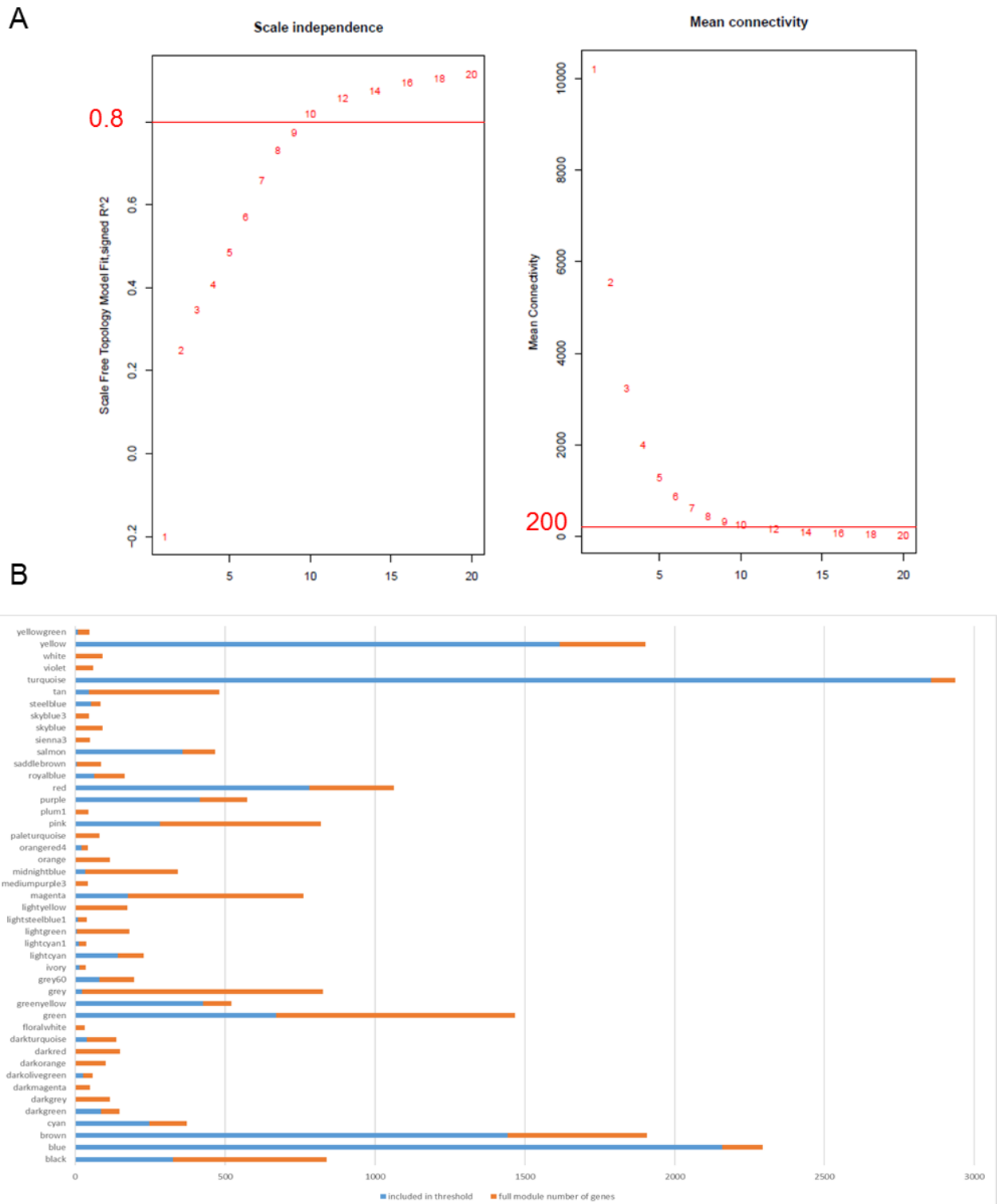
- Zhang, et al. 2018. “ Sweet Sorghum Originated through Selection of Dry , a Plant-Specific NAC Transcription Factor Gene .” *The Plant Cell* 30 (10): 2286–2307. <https://doi.org/10.1105/tpc.18.00313>.
- Zhang, W., and R. Wu. 1988. “Efficient Regeneration of Transgenic Plants from Rice Protoplasts and Correctly Regulated Expression of the Foreign Gene in the Plants.” *Theoretical and Applied Genetics* 76 (6): 835–40. <https://doi.org/10.1007/BF00273668>.
- Zhang, Yang, Jianbin Su, Shan Duan, Ying Ao, Jinran Dai, Jun Liu, Peng Wang, et al. 2011. “A Highly Efficient Rice Green Tissue Protoplast System for Transient Gene Expression and Studying Light/Chloroplast-Related Processes.” *Plant Methods* 7 (1): 30. <https://doi.org/10.1186/1746-4811-7-30>.
- Zhao, Jian, and Richard A. Dixon. 2009. “MATE Transporters Facilitate Vacuolar Uptake of Epicatechin 3'- O -Glucoside for Proanthocyanidin Biosynthesis in Medicago Truncatula and Arabidopsis.” *The Plant Cell* 21 (8): 2323–40. <https://doi.org/10.1105/tpc.109.067819>.
- Zhao, Kangmei, and Laura E Bartley. 2014. “Comparative Genomic Analysis of the R2R3 MYB Secondary Cell Wall Regulators of Arabidopsis, Poplar, Rice, Maize, and Switchgrass.” *BMC Plant Biology* 14 (1): 135. <https://doi.org/10.1186/1471-2229-14-135>.
- Zhao, Qiao, and Richard A. Dixon. 2011. “Transcriptional Networks for Lignin Biosynthesis: More Complex than We Thought?” *Trends in Plant Science* 16 (4): 227–33. <https://doi.org/10.1016/j.tplants.2010.12.005>.
- Zhao, Wei, Peter Langfelder, Tova Fuller, Jun Dong, Ai Li, and Steve Hovarth. 2010. “Weighted Gene Coexpression Network Analysis: State of the Art.” *Journal of Biopharmaceutical Statistics* 20 (2): 281–300. <https://doi.org/10.1080/10543400903572753>.
- Zhao, Zuo-yu, Tishu Cai, Laura Tagliani, Mike Miller, Ning Wang, Hong Pang, Sheryl Schroeder, Dave Hondred, Jon Seltzer, and Dortie Pierce. 2000. “Agrobacterium -Mediated Sorghum Transformation,” no. 1991: 789–98.
- Zheng, Xing-wei, Deng-xia Yi, Lin-hui Shao, and Cong Li. 2017. “In Silico Genome-Wide Identification, Phylogeny and Expression Analysis of the R2R3-MYB Gene Family in Medicago Truncatula.” *Journal of Integrative Agriculture* 16 (7): 1576–91. [https://doi.org/10.1016/S2095-3119\(16\)61521-6](https://doi.org/10.1016/S2095-3119(16)61521-6).
- Zhong, Heng, Wenling Wang, and Mariam Sticklen. 1998. “In Vitro Morphogenesis of Sorghum Bicolor (L.) Moench: Efficient Plant Regeneration from Shoot Apices.” *Journal of Plant Physiology* 153 (5–6): 719–26. [https://doi.org/10.1016/S0176-1617\(98\)80226-5](https://doi.org/10.1016/S0176-1617(98)80226-5).
- Zhong, R., C. Lee, J. Zhou, R. L. McCarthy, and Z.-H. Ye. 2008. “A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in Arabidopsis.” *The Plant Cell Online* 20 (10): 2763–82. <https://doi.org/10.1105/tpc.108.061325>.
- Zhong, R., E. A. Richardson, and Z.-H. Ye. 2007. “The MYB46 Transcription Factor Is a Direct Target of SND1 and Regulates Secondary Wall Biosynthesis in Arabidopsis.” *The Plant Cell*

- Online 19 (9): 2776–92. <https://doi.org/10.1105/tpc.107.053678>.
- Zhong, Ruiqin, David H. Burk, W. Herbert Morrison, and Zheng-Hua Ye. 2002. “A Kinesin-Like Protein Is Essential for Oriented Deposition of Cellulose Microfibrils and Cell Wall Strength.” *The Plant Cell* 14 (12): 3101–17. <https://doi.org/10.1105/tpc.005801>.
- Zhong, Ruiqin, Taku Demura, and Zheng-Hua Z.-H. Ye. 2006. “SND1, a NAC Domain Transcription Factor, Is a Key Regulator of Secondary Wall Synthesis in Fibers of Arabidopsis.” *The Plant Cell Online* 18 (11): 3158–70. <https://doi.org/10.1105/tpc.106.047399>.
- Zhong, Ruiqin, Chanhui Lee, Ryan L. McCarthy, Cromwell K. Reeves, Evan G. Jones, and Zheng Hua Ye. 2011. “Transcriptional Activation of Secondary Wall Biosynthesis by Rice and Maize NAC and MYB Transcription Factors.” *Plant and Cell Physiology* 52 (10): 1856–71. <https://doi.org/10.1093/pcp/pcr123>.
- Zhong, Ruiqin, Chanhui Lee, and Zheng Hua Ye. 2010a. “Global Analysis of Direct Targets of Secondary Wall NAC Master Switches in Arabidopsis.” *Molecular Plant* 3 (6): 1087–1103. <https://doi.org/10.1093/mp/ssq062>.
- Zhong, Ruiqin, Chanhui Lee, and Zheng-Hua Ye. 2010b. “Evolutionary Conservation of the Transcriptional Network Regulating Secondary Cell Wall Biosynthesis.” *Trends in Plant Science* 15 (11): 625–32. <https://doi.org/10.1016/j.tplants.2010.08.007>.
- Zhong, Ruiqin, Ryan L. McCarthy, Marziyeh Haghighat, and Zheng Hua Ye. 2013. “The Poplar MYB Master Switches Bind to the SMRE Site and Activate the Secondary Wall Biosynthetic Program during Wood Formation.” *PLoS ONE* 8 (7): 18–21. <https://doi.org/10.1371/journal.pone.0069219>.
- Zhong, Ruiqin, Elizabeth A. Richardson, and Zheng Hua Ye. 2007. “Two NAC Domain Transcription Factors, SND1 and NST1, Function Redundantly in Regulation of Secondary Wall Synthesis in Fibers of Arabidopsis.” *Planta* 225 (6): 1603–11. <https://doi.org/10.1007/s00425-007-0498-y>.
- Zhong, Ruiqin, and Zheng Hua Ye. 2015a. “Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation.” *Plant and Cell Physiology* 56 (2): 195–214. <https://doi.org/10.1093/pcp/pcu140>.
- Zhong, Ruiqin, and Zheng-Hua Ye. 2015b. “The Arabidopsis NAC Transcription Factor NST2 Functions Together with SND1 and NST1 to Regulate Secondary Wall Biosynthesis in Fibers of Inflorescence Stems.” *Plant Signaling & Behavior* 10 (2): 1–7. <https://doi.org/10.4161/15592324.2014.989746>.
- Zhong, Ruiqin, and Zheng-Hua Ye. 2012. “MYB46 and MYB83 Bind to the SMRE Sites and Directly Activate a Suite of Transcription Factors and Secondary Wall Biosynthetic Genes.” *Plant and Cell Physiology* 53 (2): 368–80. <https://doi.org/10.1093/pcp/pcr185>.
- Zhong, Ruiqin, and Zheng Hua Ye. 2007. “Regulation of Cell Wall Biosynthesis.” *Current Opinion in Plant Biology* 10 (6): 564–72. <https://doi.org/10.1016/j.pbi.2007.09.001>.
- Zhong, Ruiqin, Youxi Yuan, John J. Spiekerman, Joshua T. Guley, Janefrances C. Egbosiuba, and

- Zheng Hua Ye. 2015. "Functional Characterization of NAC and MYB Transcription Factors Involved in Regulation of Biomass Production in Switchgrass (*Panicum Virgatum*).” *PLoS ONE* 10 (8): 1–24. <https://doi.org/10.1371/journal.pone.0134611>.
- Zhou, Huanbin, Bo Liu, Donald P. Weeks, Martin H. Spalding, and Bing Yang. 2014. "Large Chromosomal Deletions and Heritable Small Genetic Changes Induced by CRISPR/Cas9 in Rice.” *Nucleic Acids Research* 42 (17): 10903–14. <https://doi.org/10.1093/nar/gku806>.
- Zhou, J., C. Lee, R. Zhong, and Z.-H. Ye. 2009. "MYB58 and MYB63 Are Transcriptional Activators of the Lignin Biosynthetic Pathway during Secondary Cell Wall Formation in Arabidopsis.” *The Plant Cell Online* 21 (1): 248–66. <https://doi.org/10.1105/tpc.108.063321>.
- Zhou, Jianli, Ruiqin Zhong, and Zheng Hua Ye. 2014. "Arabidopsis NAC Domain Proteins, VND1 to VND5, Are Transcriptional Regulators of Secondary Wall Biosynthesis in Vessels.” *PLoS ONE* 9 (8). <https://doi.org/10.1371/journal.pone.0105726>.
- Zhou, Xiang, and Matthew Stephens. 2012. "Genome-Wide Efficient Mixed-Model Analysis for Association Studies.” *Nature Genetics* 44 (7): 821–24. <https://doi.org/10.1038/ng.2310>.
- Zhou, Xiaofan, Xing-Xing Shen, Chris Todd Hittinger, and Antonis Rokas. 2018. "Evaluating Fast Maximum Likelihood-Based Phylogenetic Programs Using Empirical Phylogenomic Data Sets.” *Molecular Biology and Evolution* 35 (2): 486–503. <https://doi.org/10.1093/molbev/msx302>.
- Zhu, Chengsong, Michael Gore, Edward S. Buckler, and Jianming Yu. 2008. "Status and Prospects of Association Mapping in Plants.” *The Plant Genome Journal* 1 (1): 5. <https://doi.org/10.3835/plantgenome2008.02.0089>.
- Zhu, Gengrui, Guanxing Chen, Jiantang Zhu, Yan Zhu, Xiaobing Lu, Xiaohui Li, Yingkao Hu, and Yueming Yan. 2015. "Molecular Characterization and Expression Profiling of NAC Transcription Factors in *Brachypodium distachyon* L.” *PLoS ONE* 10 (10): 1–23. <https://doi.org/10.1371/journal.pone.0139794>.
- Zhu, Jinming, Sophie Alvarez, Ellen L. Marsh, Mary E. LeNoble, In-Jeong Cho, Mayandi Sivaguru, Sixue Chen, et al. 2007. "Cell Wall Proteome in the Maize Primary Root Elongation Zone. II. Region-Specific Changes in Water Soluble and Lightly Ionically Bound Proteins under Water Deficit.” *Plant Physiology* 145 (4): 1533–48. <https://doi.org/10.1104/pp.107.107250>.

Annexes

Annexe SII-1: (A) Free topological index de WGCNA afin de choisir les pouvoir en utilisant la correspondance avec les modèles de scale independence et de mean connectivity. Les niveaux de seuils (0,8 et 200) sont définis comme recommandés dans la FAQ de WGCNA (au-dessus de 0.8 et en dessous des centaines) afin de conserver le pouvoir aussi faible que possible. Pour nos analyses un pouvoir de 12 a été utilisé. (B) Répartition des gènes dans les modules identifiés par l'analyse de réseaux de co-expression. Les nombre de gènes au-dessus et en-dessous du seuil de corrélation de 0.82 (respectivement en blue et orange) sont fournis.



Annexe SII-2: GO-terms liés aux parois agrégés par mot-clé. Afin de faciliter l'analyse d'enrichissement en GO-term, une agrégation des GO-terms liés aux parois a été réalisée.

Mot-clé	GO-term ID	Processus Biologiques	Annoté ¹
SCW	GO:0009834	plant-type secondary cell wall biogenesis	112
	GO:1901347	negative regulation of secondary cell wall biogenesis	3
	GO:1901348	positive regulation of secondary cell wall biogenesis	4
	GO:2000652	regulation of secondary cell wall biogenesis	51
CW	GO:0052386	cell wall thickening	74
	GO:0016998	cell wall macromolecule catabolic process	21
	GO:0042546	cell wall biogenesis	263
	GO:0042547	cell wall modification involved in multidimensional cell growth	9
	GO:0009831	plant-type cell wall modification involved in multidimensional cell growth	5
	GO:0009664	plant-type cell wall organization	186
	GO:0042545	cell wall modification	177
	GO:0009828	plant-type cell wall loosening	22
	GO:0009830	cell wall modification involved in abscission	4
	GO:0071668	plant-type cell wall assembly	33
	GO:0006075	(1->3)-beta-D-glucan biosynthetic process	15
Lignin	GO:0046274	lignin catabolic process	43
	GO:1901430	positive regulation of syringal lignin biosynthetic process	2
	GO:0009808	lignin metabolic process	132
	GO:0009809	lignin biosynthetic process	72
Xylan	GO:0045492	xylan biosynthetic process	44
	GO:1990937	xylan acetylation	3
	GO:0015790	UDP-xylose transmembrane transport	5
	GO:0042732	D-xylose metabolic process	6
	GO:0010417	glucuronoxylan biosynthetic process	30
	GO:0045493	xylan catabolic process	20
	GO:0010411	xyloglucan metabolic process	90
Cellulose	GO:0030244	cellulose biosynthetic process	80
	GO:0052324	plant-type cell wall cellulose biosynthetic process	13
	GO:0010215	cellulose microfibril organization	23
	GO:0052541	plant-type cell wall cellulose metabolic process	20
Pectin	GO:0010400	rhamnogalacturonan I side chain metabolic process	2
	GO:0045489	pectin biosynthetic process	25
	GO:0010306	rhamnogalacturonan II biosynthetic process	5
	GO:0010246	rhamnogalacturonan I biosynthetic process	3
	GO:0010253	UDP-rhamnose biosynthetic process	3
	GO:1902066	regulation of cell wall pectin metabolic process	13
	GO:0045490	pectin catabolic process	7

	GO:0052325	cell wall pectin biosynthetic process	13
	GO:0048354	mucilage biosynthetic process involved in seed coat development	64
	GO:0033481	galacturonate biosynthetic process	4
	GO:0019567	arabinose biosynthetic process	5
	GO:0010405	arabinogalactan protein metabolic process	14
Phenylpropanoid pathway	GO:0033587	shikimate biosynthetic process	2
	GO:0019632	shikimate metabolic process	7
	GO:0009699	phenylpropanoid biosynthetic process	137
	GO:0009805	coumarin biosynthetic process	2
	GO:0009094	L-phenylalanine biosynthetic process	6
	GO:0006559	L-phenylalanine catabolic process	15
	GO:0080088	spermidine hydroxycinnamate conjugate biosynthetic process	1
Flavonoid	GO:0009812	flavonoid metabolic process	206
	GO:0031539	positive regulation of anthocyanin metabolic process	50
	GO:1900384	regulation of flavonol biosynthetic process	11
	GO:0010023	proanthocyanidin biosynthetic process	6
	GO:0009715	chalcone biosynthetic process	9
Vessels development	GO:0090058	metaxylem development	8
	GO:0010089	xylem development	54
	GO:0010222	stem vascular tissue pattern formation	7
Wax	GO:0010345	suberin biosynthetic process	33
	GO:0010143	cutin biosynthetic process	68

¹ Nombre de genes annotés avec ledit GO ID dans notre base de données de GO-terms

Annexe SIII-1 : Milieux utilisés pour les essais de transformation par biolistique

Components	Stock Conc	Callus Induction Media (CIM)	Osmotic Medium	Regeneration Medium	Rooting Medium
MS with Gamborg vit	-	4.44 g/L	4.44 g/L	4.44 g/L	4.44 g/L
Sucrose	-	30 g/L	30 g/L	30 g/L	30 g/L
Agar	-	8 g/L	8 g/L	8 g/L	8 g/L
L-proline	-	1 g/L	-	-	-
L-asparagine	-	1 g/L	-	-	-
Potassium dihydrogenphosphate (KH ₂ PO ₄)	-	1 g/L	-	-	-
D-Manitol			36.4 g/L (0.2M)	-	-
D-Sorbitol			36.4 g/L (0.2M)	-	-
Copper Sulfate (CuSO ₄)*	5 mM	200 uL/L (1 uM)	-	200 uL (1 uM)	-
2,4-Dichlorophenoxyacetic acid (2,4-D)	2 mg/L	500 uL/L (1 mg/L)	-	-	-
6-Benzylaminopurine (BAP)*	2 mg/L	-	-	500 uL/L (1 mg/L)	-
Indole-3-acetic acid (IAA)*	2 mg/L	-	-	500 uL/L (1 mg/L)	500 uL/L (1 mg/L)
Indole-3-butyric acid (IBA)*	2 mg/L	-	-	-	500 uL/L (1 mg/L)
1-Naphthaleneacetic acid (NAA)*	2 mg/L	-	-	-	500 uL/L (1 mg/L)
Geneticin (G418)*	50 mg/mL	-	-	Sel: 600 uL/L	Sel: 600 uL/L

*Doit être ajouté après autoclave

Annexe SIII-2: Milieux de culture in vitro utilisés pour la transformation par Agrobacterium

Components (stock concentrations)	Inoculation medium (liquid)^a	Cocultivation medium^a	Callus induction medium^a	Selection medium I^a	Selection medium II^a	Selection medium III^a	Regeneration medium^a	Rooting medium^a
Murashige and Skoog basal salts	4.4 g	4.4 g	4.4 g	4.4 g	4.4 g	4.4 g	4.4 g	2.2 g
Nicotinic acid (5 mg/mL)	100 µL	100 µL	100 µL	100 µL	100 µL	100 µL	100 µL	50 µL
Pyridoxin HCl (5 mg/mL)	100 µL	100 µL	100 µL	100 µL	100 µL	100 µL	100 µL	50 µL
Thiamine HCl (10 mg/mL)	100 µL	100 µL	100 µL	100 µL	100 µL	100 µL	100 µL	50 µL
Myo-inositol	0.1 g	0.1 g	0.1 g	0.1 g	0.1 g	0.1 g	0.1 g	0.05 g
2,4-Dichlorophenoxy acetic acid (2,4-D) (0.5 mg/mL)	4 mL	4 mL	4 mL	4 mL	4 mL	4 mL	–	–
Sucrose	68.5 g	20 g	30 g	30 g	20 g	15 g	–	–
Glucose	36 g	10 g	–	–	–	–	–	–
Mannose	–	–	–	20 g	15 g	10 g	10 g	10 g
2-(<i>N</i> -morpholino) ethane-sulfonic acid (MES)	0.6 g	0.6 g	–	–	–	–	–	–
Proline	–	0.7 g	–	–	–	–	0.7 g	–

Indole acetic acid (IAA) (1 mg/mL)	–	–	–	–	–	–	200 µL	200 µL
Kinetin (1 mg/mL)	–	–	–	–	–	–	500 µL	–
1-Naphthaleneacetic acid (NAA) (1 mg/mL)	–	–	–	–	–	–	–	200 µL
Indole-3-butyric acid (IBA) (1 mg/mL)	–	–	–	–	–	–	–	200 µL
MS Micronutrient solution	100 mL	100 mL	100 mL	100 mL	100 mL	100 mL	–	–
pH ^b	5.6	5.2	5.8	5.8	5.8	5.8	5.8	5.8
Phytigel	2.0 g	2.0 g	2.0 g	2.0 g	2.0 g	2.0 g	2.0 g	2.0 g
Autoclave for	22 min	22 min	22 min	22 min	22 min	22 min	22 min	22 min
Acetosyringone (200 mM)	500 µL	500 µL	–	–	–	–	–	–
Carbenicillin (100 mg/L)	–	–	1 mL	1 mL	1 mL	1 mL	1 mL	1 mL

^a Quantités indiquées pour 1L de milieu

Annexe SIII-3 : Milieux utilisés pour la production et la transformation de protoplastes

		pour 200ml de solution
Mannitol	600mM	21,86 g
Mes	2mM	0.076 g

SD (Solution Digestion)		pour 200ml de solution
CaCl ₂ , 2H ₂ O (MW 147)	7.5mM	222mg
Mannitol (MW 182.17)	600mM	21,86 g
MES buffer (MW 195,24)	10mM	390mg

W5		pour 1000ml
NaCl (MW 58.4)	154mM	8.92g
CaCl ₂ , 2H ₂ O (MW 147)	125mM	18.36g
KCl (MW 74.55)	5mM	0,372g
MES buffer (MW 195,24)	2mM	0,388g
osmolarité	634.83mosmol/L	

MMg2		pour 200ml
Mannitol	400mM	14,56g
MgCl ₂ , 6H ₂ O (MW 203.3)	15mM	0,61g
MES buffer (MW 195,24)	4mM	0,160g
Cacl ₂ 2H ₂ O	5mM	111mg
Osmolarité	464.08mosmol/L	

PEG 40%		pour 50ml
D-Mannitol	XmM	0,364g
Ca(NO ₃) ₂ ·4H ₂ O	XmM	1,18g
PEG6000 Fluka®	40%	20g

Solution enzymatique		pour 20ml de solution
Macerozyme R10 Onozuka (Duchefa 00651803)	0,75%	150mg
Cellulase RS Onozuka (Duchefa 00578704)	1,5%	300mg
BSA 10%	0,1%	200ul solution 100X
QSP Solution de digestion (SD)		20ml